Constraints and tradeoffs: Toward a predictive, mechanism-based understanding of ecological communities

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Adam Thomas Clark

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Dr. David Tilman

August, 2017

Acknowledgements:

Funding and resources:

During my graduate studies, I have been very fortunate to receive generous funding from a number of sources. These include living stipends through the University of Minnesota Graduate Excellence Fellowship, the National Science Foundation (NSF) Graduate Research Fellowship Program (base award #00006595), and the International Balzan Prize Foundation (awarded to D. Tilman). I have also received research and travel funds from numerous awards to D. Tilman, as well as the Florence Rothman Fellowship, the University of Minnesota Sigerfoos Fellowship, the University of Minnesota Department of Ecology, Evolution, and Behavior Research and Summer Fellowships, and the Graduate and Professional Student Association Travel Grant. Computing resources were provided by XSEDE through the NSF, and the University of Minnesota Supercomputing Institute. Lastly, I have benefited greatly from projects funded through the NSF's Long Term Ecological Research Network (DEB-8114302, DEB-8811884, DEB-9411972, DEB-0080382, DEB-0620652, and DEB-1234162), the University of Minnesota, and the Cedar Creek Ecosystem Science Reserve.

Direct Contributions:

I am enormously grateful for advice, mentoring, and support from my advisor, Dave Tilman, and my committee, Elizabeth Borer, Clarence Lehman, and Claudia Neuhauser. I would also like to thank J. Krueger, T. Mielke, and K. Worm for their assistance in planning and implementing fieldwork at Cedar Creek, as well as J. Anderson, D. Bahauddin, P. Barnes, S. Barrott, J. Miller, C. Potter, M. Saxhaug, and M. Spivey for help and support during my field seasons. In addition to my advisor and committee, I would like to thank many people for advice and feedback on early drafts of the chapters presented here, including: M. Burgess, C. Canham, M. Clark, W. Clark, J. Cowles, N. Eisenhauer, C. Farrior, G. Fury, B. Haegeman, F. Isbell, K. Kimmel, S. Loberg, M. Loreau, S. Pacala, D. Renard, M. Thakur, K. Thompson, D. Williams, P. Wragg, and several anonymous reviewers. Additionally, I am grateful to S. Weisberg for advice on statistical methods. Specific contributions are listed within each chapter.

Subsequent drafts of the chapters presented here are intended for later submission to academic journals. A revised and expanded version of Chapter One will be submitted including J. Knops and D. Tilman as co-authors. A revised and expanded version of Chapter Two will be submitted with C. Neuhauser as a co-author. A revised and expanded version of Chapter Three will be submitted with C. Neuhauser and D. Tilman as co-authors. Lastly, a revised and expanded version of Chapter Four is currently under review, with C. Lehman and D. Tilman as co-authors.

Indirect contributions:

I would like to thank several groups of people for their generous mentoring and support over the past six years. These include: *(1)* many researchers at University of Minnesota, who substantially expanded my knowledge and experience, including S. Binder, H. Flores, Y. Hautier, F. Isbell, B. Janke, E. Lind, K. Mueller, W. Pearse, A. Schweiger, and D. Stanton; *(2)* the Tilman lab, including P. Hawthorne, R. Putnam, M. Kosmala, M. Burgess, P. Wragg, J. Cowles, M. Clark, G. Furey, K. Kimmell, and E. Strombom; *(3)* the graduate student body at University of Minnesota, and in particular, the other students in the 2011 cohort in the Department of Ecology, Evolution, and Behavior; *(4)* my students and interns in various courses and field projects over the past six years, both for teaching me to be a better teacher, and helping me pick up some important extra tidbits about ecology along the way; *(5)* the staff in the Department of Ecology, Evolution, and Behavior at the University of Minnesota, and in particular, L. Wiggins, for helping keep me on track; *(6)* the faculty, staff, and students at the University of California, Santa Barbara, for taking me in during the winter terms that I spend in southern California; *(7)* The Sugihara lab, including G. Sugihara, E. Deyle, C. Perretti, and H. Ye, for generously introducing me to their exciting new analytical methods; *(8)* the scientists at the Smithsonian Tropical Research Institute, including M. Detto, H. Muller-Landau, S. Schnitzer, and J. Wright, for introducing me to the world of tropical ecology; and *(9)* my friends and family, including the Clarks, the Thompsons, and the various generations of roommates and grill buddies at 2312 Priscilla Street, for their emotional support and kindness.

Dedication:

To my teachers and mentors, in particular Dave Tilman, Brian Farrell, George Buckley, and Nancy Samaria, for showing me how to be a scientist.

To my family, Anni, Bill, and Graham, for their support and love.

And to Katie, for inspiring me to do good, both inside and outside of ecology.

Abstract:

It is easy to forget how far ecology has come in a very short time. Less than two decades ago, it was unclear whether predictive models of species-level dynamics in diverse ecological communities would ever be possible. Today, an abundance of methods can accurately forecast these dynamics, driven by explosive growth in the availability of data and modern analytical tools. However, most of these methods rely on matching patterns from historical dynamics to current trends. Thus, while predictions have become much easier, understanding why behavior occurs – and extrapolating predictions to novel circumstances – remains elusive.

Here, we apply theoretical insights from tradeoffs to better understand how species in ecological communities assemble and coexist. Tradeoffs describe physiological and ecological constraints that limit the traits and roles of individual organisms. These constraints therefore contain substantial information about species ecological and evolutionary histories, and how they are likely to interact with one another and their environments. We show that information contained in tradeoffs can be used to identify important mechanisms governing community dynamics, and to constrain viable parameter space in otherwise intractable models. These methods could substantially improve mechanism-based predictions in diverse communities, resulting in better understanding of how these complex systems function, and better extrapolations of predictions under novel circumstances.

Table of Contents:

List of Tables:

List of Figures:

Introduction:

By many accounts, a primary goal of ecology – insofar as any field as broad as ecology can have a primary goal – is to understand and explain the distribution and abundance of species (Lehman, Loberg & Clark 2017). Though many pathways have yielded insights that help address this goal, the route that intrigues me the most is ecological theory. In order to properly develop an intuitive understanding of the nuance and unpredictability of living things, an ecologist clearly needs a solid empirical background, with all the mud, lost blood, and chaos that comes with fieldwork. But in order to piece together all the small things that we see in the field into a larger-scale picture of the meaningful "repeated patterns" of the ecological world, familiarity with ecological theory is just as important.

Following a relatively common trajectory, it was only late in my education that I developed a proper appreciation for theory. Studying ant taxonomy and biogeography as an undergraduate, I was privileged enough to be able to travel across the United States and the Caribbean, collecting and documenting species distributions. While great fun and full of adventure, I must admit that these expeditions likely did very little to push back the boundaries of science. The thing that finally drew me into theoretical research – and, hopefully, a more productive career – was a book given to me by my father near the end of my time as an undergraduate: *Geographical Ecology*, written by Robert MacArthur (1972). Beyond the excitement of reading through my father's notes in the margin – written during his own time as a graduate student – I found myself particularly drawn to a proof in the book's appendix. There, MacArthur derived a simple competition model, which he argued could be parameterized using field measurements of species consumption rates for various resources. This was one of the first indications to me that predictive, mechanism-based theories – that is, theories that were intended not only to match patterns, but explain why those patterns existed – could be directly linked to empirical observations from the real world.

Without intending it, I have found in retrospect that MacArthur's resource competition model, or other models that are arguably its intellectual offspring, have managed to permeate every chapter of this dissertation. In this, my work is very much not

alone. Remarkably similar models have been successfully applied to mimic population dynamics of animals (Wilbur 1972), plants (Wedin & Tilman 1993), protists (Carrara *et al.* 2015), single celled algae (Tilman 1976), fungi (Gause 1932), bacteria (Coyte, Schluter & Foster 2015), and viruses (May & Anderson 1987), to name a few examples. These species span the breadth of the biological kingdoms, and are separated by more than two billion years of genetic divergence. Yet despite vast differences in cellular structure (or, indeed, the existence of cells in the first place), metabolic chemistry, physiology, resource requirements, and countless other features, the ways in which populations and communities of these species grow and interact are startlingly alike. Paraphrasing one of my committee members, Clarence Lehman, it would seem that ecological interactions, perhaps even more than our shared genetic code, are the things that unite all of us across the many branches of the tree of life.

This dissertation focuses on tradeoffs, which are the best explanation I know of for why these similarities exist across so many types of organisms. Broadly, tradeoffs describe constraints among alternative strategies that cause investments in one strategy to unavoidably come at a cost to the others. Tradeoffs are therefore important both in structuring how species evolve in response to ecological and environmental changes (Stearns 1989), and in determining the combinations of species that can coexist in any given system (Tilman 1982; Chesson 2000; Chase & Leibold 2003). There is substantial evidence that all organisms are subject to the same underlying tradeoffs, and that their responses to one another and to their environments are therefore regulated by similar forces (Tilman 1990, 2011). However, beyond this broad-scale definition and a general acknowledgement of their importance, different disciplines identify tradeoffs differently.

Twentieth century views often invoked tradeoffs as general explanations for deviations between observed and theoretically "optimal" trait values – for example, that the precise thickness of a brachiopod's shell might indicate tradeoffs between strength and resource investment (Gould & Lewontin 1979). More recent evolutionary frameworks tend to define tradeoffs in terms of fitness costs associated with changes in specific traits (Roff 1992; Roff & Fairbairn 2007), such as increased mortality risk associated with higher reproductive rates (Reznick 1985). Ecologists, on the other hand,

usually describe tradeoffs either in terms of empirical patterns, or specific theoretical models. For example in plant physiology, broadly conserved relationships among functional traits are often referred to as tradeoffs (Wright *et al.* 2004; Reich 2014), whereas in most models of ecological competition, tradeoffs describe the relationships among parameters that are required for coexistence (Tilman 1982, 1990, 2011; Chesson 2000). Perhaps the strictest definition occurs in frameworks such as the "metabolic theory of ecology", where tradeoffs arise from carefully specified low-level assumptions about species physical structure and the flow of energy and resources (Brown *et al.* 2004).

As the title of this dissertation suggests, it is intended largely as a continuation of work started more than two decades ago by my advisor, in his article *Constraints and Tradeoffs: Toward a Predictive Theory of Competition and Succession* (Tilman 1990). In that publication, and in several subsequent works (Tilman 2004, 2011), tradeoffs have been defined as the joint effect of both physiological constraints, which limit the range of traits that a single organism can theoretically possess, and of ecological constraints, which limit the range of traits that allow a species to persist within a particular ecosystem. Critically, this definition implies that tradeoffs include information about how species are likely to interact, respond to novel circumstances, and evolve. The overarching goal of this dissertation is to develop methods that help expand the range of conditions across which accurate predictions of ecological dynamics can be made, and to improve understanding of diverse ecosystems across large spatial and temporal scales. In the following chapters, we work to explain in somewhat greater detail what kinds of useful information are likely to be contained in tradeoffs, identify how to extract and use this information, and demonstrate that without this information, it may well be impossible to develop a predictive understanding of diverse ecological systems.

Chapter descriptions:

In Chapter One, *Determinism and stochasticity during 88 years of grassland succession*, we analyze plant community dynamics in abandoned agricultural fields at the Cedar Creek Ecosystem Science Reserve in Minnesota, USA (Cedar Creek). This chapter combines classic pattern-matching methods with results from on-site experiments to

identify potential drivers of observed successional dynamics. Using a series of regressions fit to observed abundance trends of major functional groups, we parameterize a metapopulation model and estimate how community dynamics are influenced by environment and by interspecific competition. In accordance with results from previous studies, we find that that successional trajectories of major species groups at Cedar Creek are relatively consistent and predictable. However, we also show that these methods perform poorly for rarer species groups, often failing to explain as much as 90% of observed variation.

In Chapter Two, *Harnessing uncertainty to Approximate mechanistic models*, we offer a potential theoretical explanation for the poor predictive power observed for rare species in the previous chapter. Using a modified version of MacArthur's resource competition model, we track how uncertainty propagates through phenomenological estimates of species interaction strengths. We show that underlying mechanisms that structure species interactions – for example, the specific type and number of resources for which species compete – strongly constrain model parameter space in ways that are not easily captured by naïve, pattern-matching methods. We also demonstrate how a hypothesized mechanistic "backbone" can be used to develop a hybrid model. This hybrid approach captures some of the underlying structure of system dynamics, which can help improve predictions, and identify the kinds of mechanisms that are responsible for observed dynamics.

In Chapter Three, *Diversity, high dimensional tradeoffs, and trait evolution buffer effects of reduced resource limitation on coexistence*, we show how specific consideration of physiological and ecological tradeoffs can yield more precise insight into how diverse communities of species interact and evolve. Using a series of models adapted from those in the previous chapter, we show that ecological constraints can cause trait divergence among species that share the same physiological tradeoffs. More importantly for the purposes of this dissertation, we also show that ecological tradeoffs can rapidly drive communities towards low-dimensional trait spaces, defined by the factors that are most important for local coexistence. This result implies that even if

physiological tradeoffs are high-dimensional, lower-dimensional relationships among traits may still emerge because of ecological forces.

Lastly, in Chapter Four, *Tradeoff-based mechanisms predict coexistence, productivity, and species abundances in grassland plant communities*, we apply the insight from the previous chapter to experimental plant communities at Cedar Creek. Using data from monoculture plots, we identify a tradeoff among three species-level traits, and use these to parameterize an "empirical tradeoff surface". We then parameterize a mechanism-based resource competition model with these monoculture traits, and use it to predict species abundances in multi-species mixtures. Importantly, we find that when traits are "snapped" to the nearest point on the empirical tradeoff surface, predictions improve substantially, and we are able to explain roughly 50% of variation in species-level biomass, and 60% of variation in total community biomass. These results suggest that traits and mechanisms included in the model and empirical tradeoff are important determinants of local coexistence and community structure.

Future directions:

A potential limitation of the methods that we present in this dissertation is that they are primarily focused on resource competition, and on experimental systems at Cedar Creek. Nevertheless, theoretical results from these chapters suggest that our findings should be broadly generalizable across many other sites and systems. The next logical step is to attempt to apply these approaches in other locations and for other types of species interactions. This is precisely the goal of the post-doctoral position that I will be starting in October, 2017, working with Stan Harpole and Helmut Hillebrand at the German Centre for Integrative Biodiversity Research.

Adam Clark

St. Paul, August 2017

Chapter One:

Determinism and stochasticity during 88 years of grassland succession:
Roles of soil fertility, fire, climate, competition, dispersal, and mortality

Abstract:

"Old fields" are ecosystems that have been previously managed and subsequently abandoned, usually from agricultural use. These systems can provide natural laboratories for testing hypotheses about ecological community assembly. Furthermore, old fields, and managed land that could someday become old fields, make up around a third of global ice-free land area. Understanding how old fields develop is therefore of great theoretical and empirical importance. However, old field succession can be difficult to predict: seemingly similar fields often diverge in terms of species composition and environmental conditions.

Here, we quantify the role of five potential drivers in shaping successional dynamics: soil fertility, fire, climate, competition, and metapopulation processes (i.e. colonization and mortality). We use data from three decades of surveys in twenty-four old fields at the Cedar Creek Ecosystem Science Reserve in Minnesota, USA, spanning almost a century of succession. First, we fit regressions estimating abundance, colonization, and mortality for five major species groups as a function of soil, fire, climate, and competition. We then use coefficient estimates from these regressions to parameterize a metapopulation model, with which we test for changes in successional dynamics associated with each driver.

Abundance, colonization, and mortality rates for species groups were consistently associated with the drivers. However, even large changes in soil fertility, burning regime, and climate had modest effects on predicted successional dynamics, and minor effects on abundance late in succession. This robustness of successional dynamics appears to result from compensatory tradeoffs reflected in the metapopulation model. For example, though higher soil fertility was predicted to increase abundance for most species groups, it also increased abundances of competitors, lowered colonization rates, and increased mortality.

Successional dynamics among old fields at Cedar Creek follow largely consistent trends. Though dynamics of individual fields vary, much of this variation can be

explained by the five drivers we consider. Furthermore, stochastic processes such as weather, colonization, and mortality do not appear to be sufficiently strong to cause divergence in successional trajectories among fields with similar sets of drivers. Our results therefore suggest that observed variability in successional dynamics may indicate potentially complex, but fundamentally predictable underlying processes.

Introduction:

Succession describes the naturally occurring transitions in ecological communities and environments that follow a major disturbance. We focus on "old field" succession, which follows the abandonment of active land management, most typically beginning as ploughland or pasture (Hobbs, R. J. & Walker, L. R. 2007). Succession in old fields is often viewed as a semi-natural experiment of ecosystem assembly processes, and has inspired foundational concepts in community ecology (Clements 1916; Gleason 1926; Tansley 1935). Moreover, since the start of European colonization, 10% of all land in the contiguous United States has been farmed and abandoned (Zumkehr & Campbell 2013), and globally more than 30% of all ice-free land is under some form of management that could undergo succession upon abandonment (Ramankutty & Foley 1999; Asner *et al.* 2004). Understanding how old fields develop is therefore of great importance.

Despite their rich history and broad spatial extent, an enduring question about old fields is whether observed dynamics are driven by predictable, deterministic processes (Clements 1916; Tansley 1935; Foster & Tilman 2000; Lortie *et al.* 2004; Pickett & Cadenasso 2005) – i.e. do outcomes among similar fields inevitably follow the same successional trajectories, or do they diverge over time? This question is probably best exemplified through the debate over the utility of chronosequences (also called "chronoseries", or "space-for-time substitution"), in which multiple fields that were abandoned at different times are compared as a proxy for temporal dynamics. The assumption underlying this method is that successional dynamics are similar across all replicate fields in the chronosequence. However, depending on the system or characteristic, patterns from chronosequences sometimes match dynamics observed in individual fields (e.g. soil characteristics or abundance of common species), and

sometimes do not (e.g. species richness or abundance of rare species) (Inouye *et al.* 1987; Pickett 1989; Foster & Tilman 2000; Walker *et al.* 2010). Because of these differences in dynamics among seemingly similar fields, it remains unclear what kinds of processes lead to divergence in successional trajectories, and whether accurate predictions of long-term dynamics in old fields are possible.

Two kinds of factors could lead to divergence in successional trajectories among old fields. First, deterministic successional dynamics can be influenced by variable starting conditions. For example, differences in historical management strategies (Sanderson *et al.* 2004; Fargione *et al.* 2008), or in the intensity and duration of management (Uhl, Buschbacher & Serrao 1988; Burke, Lauenroth & Coffin 1995; Post & Kwon 2000) can lead to changes in environmental conditions or local species pools. Similarly, even small differences in local soil composition, topography, or climate can initially favour some species over others, and may lead to long-term divergence in species composition (Whisenant, Thurow & Maranz 1995; Clark & Tilman 2010; Staver, Archibald & Levin 2011).

Second, stochastic processes can lead to differences among otherwise identical old fields by random chance (Norden *et al.* 2015). Examples of potential stochastic drivers of divergence include extreme weather (Tilman & Haddi 1992; Haddad, Tilman & Knops 2002), dispersal or mortality events (Bartha *et al.* 2003; Lortie *et al.* 2004), disturbance (Lortie *et al.* 2004; Hobbs, R. J. & Walker, L. R. 2007) or different initial abundances of species (Fukami *et al.* 2005; van de Voorde, van der Putten & Bezemer 2011). Though often transient, stochasticity can lead to long-term changes either through repeated events (Noble & Slatyer 1980; Fukami & Nakajima 2011), or if initial differences are subsequently reinforced or magnified by deterministic priority effects (Fukami *et al.* 2005; Stark & Norton 2015) through processes such as plant soil feedbacks (van de Voorde *et al.* 2011; Stark & Norton 2015) or environmental regime shifts (Staver *et al.* 2011).

Here, we revisit a long-term study of old field succession at the Cedar Creek Ecosystem Science Reserve in Minnesota, USA. Vegetation dynamics in these fields have been extensively described in previous publications (Inouye *et al.* 1987; Tilman

1987; Foster & Tilman 2000), and surveys have since expanded to cover more than three decades of succession across 24 fields, that when combined form an 88 year chronosequence. Furthermore, the study is nested within a site that includes experimental manipulations of soil nutrients, burning regimes, climate variables, and species composition. Because of this combination of long-term observational and experimental data, Cedar Creek is particularly well-situated to test the effects of a wide range of processes on long-term successional dynamics. Importantly, though there are some broad trends in vegetation dynamics that are shared among chronosequences at Cedar Creek, specific trends at the level of individual fields differ (Fig. 1; also see Fig. S2-S3 in the supplement for diversity and climate trends, respectively). Thus, there should be sufficient variability in observed dynamics to effectively identify some of the major drivers of divergence.

Our primary goal is to better understand the roles of deterministic and stochastic processes in shaping old field dynamics, and to determine whether these processes preclude accurate forecasting of successional trajectories at Cedar Creek. We proceed in three steps: *(1)* we fit regression models to identify major deterministic and stochastic drivers of plant community successional dynamics at Cedar Creek; *(2)* we incorporate these drivers into a basic metapopulation model to determine how they alter plant abundance, and colonization and mortality rates over the course of succession; and *(3)* we use the metapopulation model to test whether observed variability in these drivers was sufficiently large to cause divergence among successional trajectories, and if so, whether this divergence can be attributed to deterministic differences among fields, or stochastic processes.

Though many types of species and environmental factors influence and are influenced by succession, we limit our scope to changes in abundance through time among non-woody plant species, as we have the greatest empirical evidence testing long-term dynamics for this group. Similarly, we focus on five potential drivers of successional dynamics, which have been identified as particularly influential through previous work at Cedar Creek (see Tilman et al. (2012) for a general summary): soil fertility (Tilman 1987; Isbell *et al.* 2013), burning (Axelrod & Irving 1978; Cavender-

Bares & Reich 2012), climate (Tilman & Haddi 1992; Haddad *et al.* 2002), interspecific competition (Wedin & Tilman 1993; Dybzinski & Tilman 2007), and metapopulation effects of colonization and mortality across sites (Tilman 1994).

Materials and Methods:

Study system

The natural history of old fields at Cedar Creek is described in detail by Inouye et al. (1987) and Tilman (1987). Summarized briefly: soils are well-draining, composed primarily of fine and medium grain sands (Grigal *et al.* 1974), and are strongly nitrogen limited (Tilman 1987). Total atmospheric nitrogen deposition is low at roughly 1 g $m^{-2}yr^{-1}$ (± 0.6 SD; 60-70% as wet deposition), and has been relatively consistent for at least three decades (NADP 2017). Mean annual precipitation (measured between 1963 and 2016) is roughly 77 cm (±16 SD), of which about 60% falls between April and August. Daily summer high temperatures (June-August) are around 27°C, and winter lows (December-February) are around -14°C. Overall, climate change is expected to lead to warmer conditions in the region, with dryer summers and wetter winters (Kling *et al.* 2003).

The area around Cedar Creek was first colonized by Europeans in 1856, following expatriation of indigenous people as a result of treaties signed with the Dakota and Ojibwe in 1837 (MHS 2017). The earliest recorded land clearing occurred in 1885, and extensive agricultural land use commenced primarily in the first decade of the twentieth century (Pierce 1954). Prior to this, the region contained a mix of prairie, oak savanna, deciduous forests, and wetlands (Cushing 1963). Despite significant land clearing, aerial photography beginning in 1938 suggests that several large patches of forest and savanna were never cleared or ploughed (MHAPO 2015). Though fire plays an important role in maintaining prairies and savannas at Cedar Creek (Axelrod & Irving 1978), succession does not consistently lead to rapid afforestation in the absence of fire, and some fields that have not been burned in almost a century remain largely free of woody species, despite being surrounded by savanna and deciduous forest.

Study design and sampling

The old field surveys at Cedar Creek include two types of species-level measurements: percent-cover is sampled roughly every five years in 100 plots in each of 23 of the fields ("Experiment 014"), and total aboveground biomass is sampled annually in 4 plots in each of 17 fields, including one field that is not part of E014 ("Experiment 054"). These fields have been abandoned in a staggered fashion between 1927 and 2015. Percent-cover plots were initially established in 22 fields in 1983. Since then, three fields have been removed from surveys because they were transitioned into other experiments, and four newly abandoned fields have been added. In 1988, biomass sampling began in twelve of the fields sampled for percent-cover, as well as in one additional field. The same four fields added to the percent-cover surveys were also added to the biomass surveys. See Appendix A for specific details, and Fig. S1 for a map.

For both types of sampling, plots in each field are divided into four transects, each 25 m apart. Percent-cover plots are 0.5 × 1 m, with 25 plots per transect. Within a transect, plots are positioned every 1.5 m. Percent-cover of bare ground, dead leaf litter, and vegetation for each species is estimated visually, as if projected from above onto a two-dimensional surface (Inouye *et al.* 1987). Cover for all surveys is scaled to sum to 100% (i.e. relative cover), though for later surveys absolute cover was also measured. Each survey is carried out by a single group (usually 4 people), though identity of surveyors has changed since 1983. To ensure consistent measurements, surveyors train for 1-2 days using standardized cardboard cut-outs, and survey shared subsets of plots to calibrate their estimates (Tilman 1987). Species identities are confirmed in-field by a single surveyor, and voucher specimens across surveys are verified using the on-site herbarium. Biomass is measured in a 3 × 4 m plot located at the end of each transect. Measurements are taken during peak biomass (July or August) by clipping a 0.1 by 3 m strip of vegetation. Clipped vegetation is sorted to either individual species or leaf litter, and then dried and weighed. Strip locations move within the plot each year to avoid repeated clipping (Haddad *et al.* 2002). Locations for all plots are permanently marked by steel rebar.

Though woody species are relatively rare in the survey plots, a subset of fields have experienced significant woody encroachment since abandonment, mainly by *Pinus strobus* (white pine), in the northern part of Cedar Creek, and *Populus tremuloides* (quaking aspen) in the south. Heavy afforestation (i.e. substantial tree cover in all survey plots) has occurred in all four transects in one field, and in two or fewer transects in three other fields. Light afforestation (i.e. tree cover in some plots) has also occurred in two of these fields, and in at least one transect of three additional fields. Tree cover is absent from all plots in the remaining sixteen fields. Woody encroachment does not appear to be a function of field age, and many of the oldest fields in the chronosequence have not been invaded by trees. In all but four fields (the most heavily forested field and three recently abandoned fields) half of the transects in each field have received experimental burning treatments since 2006, with an achieved fire frequency of between one fire in three and one in two years.

Because nitrogen is the primary limiting soil resource at Cedar Creek (Tilman 1987), we use soil total nitrogen concentration (%N in oven-dried soil, on a mass by mass basis) as an indicator of soil fertility, which has proven a useful proxy for soil fertility in previous experiments (Wedin & Tilman 1993; Dybzinski & Tilman 2007). Soil total nitrogen concentration in mineral soil has been measured every six years since the start of the experiment in the percent-cover plots, and twice in the biomass plots (2001 and 2014). Measurements are taken using a 2.5 × 10 cm soil core from the centre of each plot after removal of litter and the O horizon, sifted using a 1 mm sieve, dried, and ground. Chemical analysis is conducted using a Carlo Urba NA 1500 elemental analyser (CE Elantech Inc., Lakewood, New Jersey; Knops & Tilman 2000). Though soil nitrogen dynamics are themselves influenced by plant community composition (Fornara & Tilman 2008), dynamics across the old fields at Cedar Creek tend to be similar, with slow accumulation as a function of field age (Fig. 2a; Knops & Tilman 2000). Thus, while we account for variable accumulation across fields in subsequent analyses, we do not directly address feedbacks between plant community composition and soil nitrogen dynamics.

To identify variability in weather and climate, we calculated annual mean daily high temperature and total precipitation for April-August for the years 1963-2016, which

captures conditions during the local growing season prior to vegetation survey or harvest. We calculated these summaries from hourly temperature and precipitation data collected at an on-site weather station. Missing records were filled in using data from nearby weather stations in Cambridge and Andover, MN, both located within about 15 km. To generate an index of historical climate conditions, we used ranged major axis regression, fit using the `lmodel2` package (Legendre 2014; version 1.7-2) in the `R` programming language (R. Development Core Team, 2016; version 3.3.2) to identify the relationship between precipitation and temperature. Following this relationship, years are generally either cool and wet, or hot and dry (Fig. 2b).

Data cleaning

We took several steps to correct and simplify data prior to analysis. First, because survey methods were not meant to collect meaningful information about trees, we excluded from the analysis transects that became heavily aforested (both before and after tree encroachment). Second, we interpolated observations of soil total nitrogen to generate estimates for each plot at the time of survey. For the percent-cover plots, we fit a linear model estimating soil total nitrogen as a function of field age (i.e. years since abandonment) with log-transformations of both variables and a random intercept nested by plot, transect, and field using the `lme4` package (Bates *et al.* 2015; version 1.1-13) in `R`. For biomass plots, where we had only two observations of soil total nitrogen per plot, we fit a linearized dynamical model following the methods of Knops and Tilman (2000). See Appendix A in the supplement for more details.

Because we identified well over 150 plant species across all surveys, we also grouped species by life history, growth form, and origin, following the methods of Inouye et al. (1987) (see Appendix A and Table S2 in the supplement for specific categorizations of each species). We further grouped these based on categories with qualitatively similar observed temporal dynamics (see Fig. 1): *(i)* perennial C4 grasses and sedges (hereafter "C4s"); *(ii)* perennial C3 grasses ("C3s"); *(iii)* annual/biennial species including both grasses and non-legume forbs ("annuals"); *(iv)* perennial non-legume forbs ("forbs"); and *(v)* legumes. Though they have similar dynamics, we separated forbs and legumes

because legumes are able to fix nitrogen, which could result in different responses to successional drivers. Importantly, these groups were present in all fields, and account for 98% of all observed cover and 99% of all aboveground biomass for non-woody species.

Identifying major drivers

We used two types of information about successional dynamics to identify important drivers: abundance of each species group, estimated from the biomass plots, and colonization or mortality rates, estimated from observed changes in presence or absence within percent-cover plots through time. For abundance, we used data from the biomass survey plots rather than the percent-cover plots because it ensured relatively independent estimates of abundance across species groups (recall percent-cover is absolute cover – thus, high abundance of one group necessarily implies low abundance of all other groups). For each species group, in plots where the summed abundance of species within that group was greater than zero, we fit regression models to observed aboveground biomass as a function of field age, soil total nitrogen concentration, burning treatment, total growing season precipitation (April-August), mean growing season high temperature, the interaction of temperature and precipitation, and the abundance of C3s and of C4s. We did not include other species groups as covariates because they were insufficiently common to produce regressions that converged consistently.

To account for nonlinear trends in the relationship between aboveground biomass and field age, we fit a general additive mixed model (GAMM) using the `mgcv` package (Wood 2011; version 1.8-17) in `R`. For each field, we fit separate thin plate splines (i.e. nonlinear smoothing functions) describing the statistical effect of age on biomass, while all other variables were modelled as simple linear effects (as in ordinary least squares). To help account for pseudoreplication and differences in sampling frequency, we included nested random intercepts for field and transect (but not plot, as this caused conversion problems). Though we would have ideally have used a regression that was able to predict species abundances entirely as a function of environmental and ecological variables, without directly including the effects of age, we found that a strong age effect remained even accounting for all other drivers, and did not want this otherwise

unexplained variation to bias our results (e.g. because of correlations between age and soil total nitrogen). All variables were scaled to zero mean and unit standard deviation prior to analysis. To meet distributional assumptions, we square root transformed abundance before scaling.

Estimates of abundance of species group i in site j at time t, $A_{i,j,t}$, were therefore fitted from the following equation:

$$A_{i,j,t} = s(age_{j,t}) + \beta_1(soilN_{j,t}) + \beta_2(burning_{j,t}) + \beta_3(precip_t) + \beta_4(temp_t) + \beta_5(precip_t \times temp_t) + \beta_6(C3_{j,t}) + \beta_7(C4_{j,t}) \quad \text{Eq. (1)}$$

where the β terms are fitted constants, $s(age_{j,t})$ is the prediction from the thin plate spline given the age of site j at time t, $soilN_{j,t}$ is the soil nitrogen concentration predicted from the interpolation function in sites j at time t, $burning_{j,t}$ is an indicator variable that is zero if a field has never been burned and one otherwise, $precip_t$ and $temp_t$ are, respectively, the observed precipitation and temperature at time t, and $C3_{j,t}$ and $C4_{j,t}$ are, respectively, the abundance of the C3 and C4 species groups in site j at time t. Note, these variables are all scaled and transformed as described above, and C3 and C4 abundance were not included in regressions predicting their own abundances. For visualization purposes, we also generated a "mean" estimate of the statistical effect of age on abundance by re-fitting regression for each species group with a single, linear age coefficient across all 88 years of chronosequence data, that were otherwise identical to the GAMMs.

For colonization and mortality rates, we used presence/absence data from the percent-cover surveys to determine rates at which species groups established or died out in survey plots (hereafter "colonization" and "mortality", respectively). Colonization was assumed when a species group was observed in a plot following a survey where the plot did not contain that group; mortality was assumed when a species group disappeared from a plot that it had previously occupied. To calculate rates, we fit generalized linear models (GLMs) using the `lme4` package. Both colonization and mortality were fit using binomial regression with a complementary log-log link function. This link function has two advantages. First, it allows inclusion of an offset term to account for different sampling intervals. Second, its coefficients can be interpreted as hazard ratios, which describe the fraction change in the likelihood of an event (see description in the legend

for Fig. 4). These regressions included all variables from the abundance regressions except for field age, because we were able to explain most of the age effect as a function of the other drivers. We also included random intercepts for field (but not plot or transect because of conversion issues). The colonization regressions also included mean within-group percent-cover in each transect as an index of local colonization pressure, and the mortality regressions also included within-group percent-cover (in the previous survey) in each plot as an index of local abundance.

Estimated annual colonization and mortality probability for species group i in site j at time t, $c_{i,j,t}$ and $m_{i,j,t}$, respectively, were therefore fitted from the following equations:

$$\text{cloglog}(c_{i,j,t}) = \beta_8(soilN_{j,t}) + \beta_9(burning_{j,t}) + \beta_{10}(precip_t) + \beta_{11}(temp_t) + \beta_{12}(precip_t \times temp_t) + \beta_{13}(C3_{j,t}) + \beta_{14}(C4_{j,t}) + \beta_{15}(Atrns_{i,j,t}) \quad \text{Eq. (2a)}$$

$$\text{cloglog}(m_{i,j,t}) = \beta_{16}(soilN_{j,t}) + \beta_{17}(burning_{j,t}) + \beta_{18}(precip_t) + \beta_{19}(temp_t) + \beta_{20}(precip_t \times temp_t) + \beta_{21}(C3_{j,t}) + \beta_{22}(C4_{j,t}) + \beta_{23}(A_{i,j,t}) \quad \text{Eq. (2b)}$$

where "cloglog" denotes the complementary log-log link function (see Appendix B.1 in the supplement), $Atrns_{i,j,t}$ is the mean abundance of group i in site j at time t across all plots in the same transect as site j (excluding site j itself), and all other terms are as described for Eq. (1).

For all abundance, colonization, and mortality regressions, we used a backwards stepwise selection procedure starting with the full model described above to remove terms that did not improve explanatory power (i.e. we set the β fitted parameter values to zero). Results are presented for these reduced models. Because potential source populations at Cedar Creek are concentrated in savannas in the southern half of the reserve, we also tested for significant effects of spatial position and autocorrelation among fields. However, after accounting for other covariates, these spatial components did not significantly improve model fit, so we do not discuss them further here. For details about the specific structure of the regressions, the link function, and the stepwise selection procedure, see Appendix B in the supplement.

Incorporating drivers into a metapopulation model

To generate predictions of successional dynamics that jointly accounted for changes in abundance, colonization, and mortality, we used the regressions outlined above to parameterize a metapopulation simulation, simplified from an earlier model describing resource competition in heterogeneous environments (Tilman 2004). In our model, we simulated annual changes in the presence, absence, and abundance of each species group across 25 "plots" (i.e. a single percent-cover sample transect). Simulations proceed by first using the fitted abundance regressions to predict abundance for each species group in plots where they were present, and then updating presence or absence using random draws based on the probability of colonization and mortality predicted from the binomial regression models from Eqs. (2a-b). All groups were assumed to be absent at year zero.

Thus, we separately tracked two types of information for each of the 25 simulated plots in this model: *(i)* presence or absence of each species group i in each plot j at time t, $p_{i,j,t}$, which was set to one if the group was present and zero if it was absent, and *(ii)* abundance of group i in plot j at time t, $A'_{i,j,t}$. Abundance was calculated following Eq. (1) as:

$$A'_{i,j,t} = A_{i,j,t} \quad \text{if } p_{i,j,t} = 1 \qquad \text{Eq. (3a)}$$

$$= 0 \quad \text{if } p_{i,j,t} = 0 \qquad \text{Eq. (3b)}$$

Presence or absence of each group in each plot changed following Eqs. (2a-b) as:

$$p_{i,j,t} = \text{rbinom}(c_{i,j,t}) \quad \text{if } p_{i,j,(t-1)} = 0 \qquad \text{Eq. (4a)}$$

$$p_{i,j,t} = 1 - \text{rbinom}(m_{i,j,t}) \quad \text{if } p_{i,j,(t-1)} = 1 \qquad \text{Eq. (4b)}$$

where rbinom(x) indicates a single draw of a random binomial variable with probability x, and returns a value of 1 if the event occurs, and zero otherwise. The model proceeded in time-steps of one year. Note that both the number of plots occupied and the abundance within occupied plots influence mean abundance calculated across all 25 simulated plots.

In this metapopulation model, deterministic changes in species group abundances depend on the relationships described in Eq. (1), including whether a plot is burned, soil total nitrogen predicted as a function of age from the linear interpolation model, and average precipitation and temperature. Stochastic variation results from colonization and

mortality events as described in Eqs. (2a-b), including, differences in the intercept of the nitrogen interpolation function (which we varied among simulations – see description of scenarios below), and random annual fluctuations in precipitation and temperature. Note that the effects of competition were therefore an emergent effect, including both deterministic responses to competitor biomass in the regression models, and stochastic presence or absence resulting from colonization and mortality events. For more details on the simulation methods, see Appendix C.1 in the supplement.

Testing for divergence

We used the metapopulation model to test for divergence using a series of scenarios. We simulated 1,000 instances of each scenario over 88 "years" of dynamics, and then tested for divergence among replicated simulations of each individual scenario, and divergence among replicated simulations of different scenarios. Within each scenario, we set all plots as either burned or unburned, and centred mean values for the intercept of the nitrogen interpolation function, precipitation, and temperature on fixed values (described below). We varied the intercept of the interpolation function between simulations by sampling from a normal distribution with standard deviation equal to the variation observed across fields. We generated inter-annual variation in weather within each simulation by sampling from the major axis regression line of the historical climate index.

We tested four main scenarios: *(i)* a "reference" scenario with no burning, intercept for the soil nitrogen interpolation centred on the mean trend observed across all fields, and precipitation and temperature centred on the mean observed in the historical climate time series; *(ii)* a "burning" scenario, which was identical to the reference except that all plots were burned; *(iii)* a "nitrogen" scenario, for which the intercept for the nitrogen interpolation function was held at two standard deviations above the observed mean (i.e. representing unusually fertile conditions); and *(iv)* a "drought" scenario, for which temperature was centred on values two standard deviations above the mean observation, corresponding to precipitation centred approximately 1.45 standard deviations below the mean (i.e. representing unusually hot, dry years). Note that this final

scenario matches the most extreme drought observed at Cedar Creek in the past 50 years (the 1988 "North American Drought"). In Appendix C.2 in the supplement, we also test a fully factorial set of burning, soil fertility, and climate scenarios. These generally matched results for the four primary scenarios, and we do not discuss them further here.

Results:

In general, species groups followed consistent successional trends across both observations and simulations (Fig. 5, top row), with early colonization by annual grasses and forbs, followed by a window of dominance by C3 grasses, followed by C4 grasses dominating. Perennial forbs and legumes followed slight time trends, but generally remained rare throughout succession. Though there were significant differences in predictions among scenarios (Fig. 5, rows 2-4), these differences were often small. Furthermore, variation among simulations within individual scenarios tended to be minor. Note that trends from raw data in Fig. 5 are for a wider range of conditions than those in the simulations, and are therefore not expected to match simulations perfectly.

Standardized effect sizes from the regression models showed that abundance, colonization, and mortality rates for species groups were strongly associated with burning, climate (i.e. combined effects of precipitation and temperature), and competition (Figs. 3d; 4a,c). These three drivers also tended to have the greatest explanatory power for both types of regressions (see partial R^2 in Fig. S4 in the supplement). Though we detected significant associations with soil total nitrogen for most species groups, estimated effect sizes and explanatory power tended to be smaller. For most groups, abundance regressions explained about 48% of total variation (i.e. model R^2), with lower explanatory power for rarer species groups (i.e. forbs and legumes). Across all groups, colonization regressions explained about 27% of variation, and mortality regressions explained about 11%.

For abundance regressions, average associations with field age generally matched observed trends from Fig. 1, with strong increases in abundance for C4s, modest increases for C3s, and strong decreases for annuals (Fig. 3d). For these groups, effect size for age was larger than that for the main drivers. Increased soil total nitrogen

concentration was associated with large increases in C4 abundance, moderate C3 increases, and moderate decreases among forbs. Burning led to strong increases in C4s, but did not significantly affect other groups. All groups were significantly and positively associated with precipitation. C3s and forbs showed reduced abundance in warmer years, while interactions of warm, wet years were associated with increases for C3s, annuals, and legumes. C3s and annuals were negatively associated with C4 abundance, while C4s, annuals, and forbs were negatively associated with C3 abundance.

In the colonization and mortality regressions, species responded to most drivers similarly (Fig. 4a,c). In comparison to the main drivers, within-group transect and plot abundance were associated with the greatest increase in colonization, or decrease in mortality, respectively. Mean colonization rates roughly matched observed dynamics, with fastest colonization among annuals, intermediate for forbs and C3s, and slowest for C4s and legumes (Fig. 4b). For all groups but C3s, increased soil total nitrogen was associated with decreased colonization and increased mortality, while burning led to increased colonization and decreased mortality. Greater precipitation and interaction of warm, wet years was generally associated with increased colonization and decreased mortality. Warmer years were associated with increased mortality among all groups, but variable responses for colonization. Increased C4 abundance tended to be associated with decreased mortality and variable responses for colonization, while increased C3 abundance was generally associated with increased mortality and decreased colonization.

Differences in metapopulation model predictions among scenarios were less pronounced than the effect sizes from the regressions. Burning led to large increases for C4s and moderate increases for forbs early in succession, but had little effect on other groups (Fig. 5 row 2). Increased nitrogen led to slight increases for C3s, and moderate decreases for annuals and forbs (Fig. 5 row 3). The only consistent response was from the drought scenario, which reduced abundance for all groups, though C4s largely recovered late in succession (Fig. 5 row 3). Interestingly, though relative abundance of species groups varied little among simulations by the end of the successional time series, they did not converge to match that of remnant grassland communities at Cedar Creek (Fig. 6).

Discussion:

Our results suggest that observed variability in successional dynamics across old fields at Cedar Creek are indicative of complex, but fundamentally consistent and predictable underlying processes. The low within-scenario variability we observe implies that differences among fields can be explained by a modest number of largely deterministic responses to drivers such as soil fertility and burning regime. Though stochastic processes such as weather, colonization, and mortality are significantly associated with species group abundance, colonization, and mortality, these processes do not appear to be sufficiently strong to cause divergence in successional trajectories among fields with similar sets of drivers. Furthermore, even large changes in burning regime, soil fertility, and climate had modest effects on predicted successional dynamics, and minor effects on relative abundance late in succession, suggesting that a single successional pathway captures much of the general trend observed across fields.

Importantly, our models include several mechanisms that have been hypothesized as potential drivers of stochastic divergence, such as year-to-year variability in temperature and precipitation, and stochastic colonization and mortality (Webb, Tracey & Williams 1972; Bartha *et al.* 2003). As such, it was entirely possible for simulations to diverge through time. For example, depending on the strength of competitive interactions and feedbacks between abundance and the colonization and mortality rates estimated in the regressions, stochastic variation caused by colonization order could have led to priority effects, thereby driving simulations towards alternate stable states, or at least alternate transient trajectories (Fukami & Nakajima 2011). Our results therefore match previous findings from old fields which suggest that both stochastic and deterministic processes leave detectable signals in successional dynamics (Myster & Pickett 1992; Norden *et al.* 2015), but differ somewhat in that we find that stochastic processes have impacts that are interpretable as shorter term noise, and do not lead to divergence.

Interpretation of models

Though they do not appear to have strong enough effects to cause divergence in successional trajectories, many of the responses to drivers detected by our regressions

and metapopulation models accord with experimental results from Cedar Creek. The response to soil fertility and burning are largely consistent with results from nitrogen fertilization experiments, which suggest that higher rates of nitrogen addition lead to increases in litter abundance, which favours C3 grasses and inhibits other groups (Tilman 1987; Clark & Tilman 2010; Isbell *et al.* 2013). Consistent with observations and experimental warming treatments at Cedar Creek, we also find decreases in abundance, colonization, and persistence among C3 grasses in warmer and dryer years, along with more rapid dominance by C4 grasses relative to other groups (Tilman & Haddi 1992; Cowles *et al.* 2016). The positive, interactive effect of warm and wet years, and the negative effect of warm and dry years observed for most groups is also consistent with general understandings of plant physiology (Lambers, Chapin & Pons 2008).

A potential explanation for why many drivers corresponded to strong estimated effect sizes in the regressions, but did not strongly influence successional dynamics or long-term outcomes in the metapopulation models, is that many of the predicted responses tended to counteract another through compensatory tradeoffs. For example, though regressions suggest that C4 abundance responds positively to soil total nitrogen concentration, this response is negated in the metapopulation simulations through a combination of changes in colonization and mortality rates and increased competition from C3 grasses. Similarly, though burning tends to increase abundance and colonization, and decrease mortality, it also increases the abundance of competitors, buffering species responses.

It is difficult to detect tradeoffs across species groups given that we consider only five groups, but there is some quantitative evidence from our fitted parameter estimates suggesting that stable successional trajectories and coexistence in our metapopulation model are indeed explained by tradeoffs. Using major axis regression to compare species mean colonization rates (Fig. 4b), mortality rates (Fig. 4d), and linear response to field age, which roughly corresponds to competitive hierarchy (Fig. 3e), we find that good colonizers tend to have a negative response to age (ρ = -0.72, p = 0.06), whereas species groups that have a positive response to age tend to have higher mortality rates (ρ = 0.95, p = 0.03). Because negative responses to age suggest that species groups are

competitively displaces (e.g. annuals), whereas positive responses suggest that species groups are able to competitively displace others (e.g. C4s), these correlations may suggest that poor competitors are able to persist because they are able to arrive before better competitors do, and that better competitors are limited by slow colonization, and potentially also high mortality early after establishment.

The general successional patterns predicted by the metapopulation model are also largely consistent with theoretical expectations for tradeoffs between dispersal and competitive ability (Tilman 1994). Annuals tend to be fast dispersers and account for a large proportion of viable seeds in remnant seedbanks, allowing them to rapidly establish in abandoned fields (Kitajima & Tilman 1996). C3 grasses tend to be intermediate dispersers and competitors, and therefore arrive after annuals, but are able to competitively displace them (Grime & Hunt 1975). Lastly, C4 grasses at Cedar Creek are generally the slowest dispersers, but are also the best nitrogen competitors, and can therefore displace C3 grasses once they arrive (Wedin & Tilman 1993; Tilman 1994). Trends among these groups also accord with results from the abundance, colonization, and mortality regressions, both in terms of competitive inhibition, and relative colonization rates. Forb and legume dynamics are somewhat less clear, potentially because they are rarer than other groups, making it more difficult to fit regressions. Alternatively, these ambiguous trends may result from the wider range of functional traits found among species within these groups, such as nitrogen fixation and deep rooting depths, which have been shown to reduce competitive inhibition by grasses (Fargione & Tilman 2005).

Convergence and divergence of successional dynamics

Our findings largely align with results from other chronosequence studies. Previous analyses of old fields at Cedar Creek were based on much shorter time series, but found similar successional dynamics across functional groups (Inouye *et al.* 1987; Foster & Tilman 2000). Similarly, many other chronosequence studies have found generally predictable patterns in the abundance and dynamics of common species groups across fields (Pickett 1989; Foster & Tilman 2000; Walker *et al.* 2010; Meiners, Pickett

& Cadenasso 2015). For example, in the Buell-Small Succession Study in New Jersey, USA, convergence among chronosequences has been detected among common species at large spatial scales, regardless of what year the field was abandoned, crop type, or farming method (Li *et al.* 2016). Correspondingly, poor predictive power from our regression models for rarer groups of species – i.e. forbs and legumes – also matches results from other chronosequence studies (Foster & Tilman 2000; Walker *et al.* 2010).

While there is substantial theoretical support for our findings suggesting a strong role of competition in structuring successional dynamics (Huston & Smith 1987; Tilman 1994; Lortie *et al.* 2004), empirical evidence is somewhat more mixed. For example, though earlier analyses of the Buell-Small Succession Study suggested that inhibition by competitively dominant grasses was common (Myster & Pickett 1992), longer-term data now suggests that community composition is more strongly structured by dispersal limitation than by competition (Li *et al.* 2015). In other regions, studies have found that competitive interactions do not become dominant until late in succession (Purschke *et al.* 2013), or that facilitation plays a more important role than competition (Zanini, Ganade & Hübel 2006). As such, though competition may be a useful predictor of successional dynamics at Cedar Creek, it is likely not equally important in all sites.

A potential factor leading to greater similarity of successional dynamics among old fields at Cedar Creek is that relatively few fields have become forested. Intriguingly, neither field age nor the proximity of potential source populations fully explains these patterns, suggesting that the dispersal and growth of trees at Cedar Creek may be exceedingly slow, and driven primarily by stochastic events (Lawson *et al.* 1999). Though insight from sites where afforestation is more prevalent might help explain why specific fields have become forested (Pickett, Cadenasso & Bartha 2001; Cook *et al.* 2005), results from some sites suggest that afforestation is, in general, a largely unpredictable process (Norden *et al.* 2015).

Further considerations

Several topics that we do not specifically consider here may provide fruitful venues for further research. First, though we only have abundance data from four

biomass plots per field, we would have preferred to include some information from the highly replicated percent-cover plots. However, because long-term surveys included only relative cover, species abundances were highly confounded, making it difficult to separate responses of different species groups from one another. Hopefully, longer time series of absolute cover, or advances in remote sensing may allow future studies to obtain data on species abundances at much larger-scales (Cavender-Bares *et al.* 2017).

Second, though we would have ideally included litter abundance as a driver, as it has strong effects on vegetation at our site (Clark & Tilman 2010; Isbell *et al.* 2013), we found that it was too strongly correlated with abundance to do so. Burning serves as something of an indicator for litter, as it leads to strong reductions in litter (Fig. 3a), though it is not a direct proxy, because burning has many other effects such as changes in the relative abundance of species (Fig. 3b).

Lastly, herbivory is known to have differential impacts on species groups in old fields (Ritchie & Tilman 1995), and has been shown to be important for maintaining grassland plant diversity at global scales (Borer *et al.* 2014). Because there are no specific manipulations of herbivore pressure in the old field surveys, we subsume its influences into the intercept terms of our regression models, which effectively assumes uniform herbivore pressure among fields. Correctly accounting for differential herbivore pressure among fields could help further improve model predictions, or could reveal additional mechanisms that lead to alternate successional trajectories (Noy-Meir 1975).

Conclusion

We find evidence that successional dynamics at Cedar Creek depend primarily on a modest number of environmental and ecological variables, and are therefore largely predictable. This result accords with previous findings suggesting that chronosequences can be helpful for describing abundance trends for common groups of species. Though it remains to be seen how frequently deterministic drivers are dominant in successional dynamics, the growing body of long-term observations such as those from the Buell-Small Succession Study (Meiners *et al.* 2015), or large-scale experiments such as the Nutrient Network (Grace *et al.* 2016), should make it increasingly feasible to identify

important drivers of successional dynamics in other systems. If variability in successional trajectories in these sites is driven by similar processes as those at Cedar Creek, then increasing spatial and temporal scales of observations should likewise lead to accurate models of successional dynamics. These models could greatly expand the tools available to land managers and conservationists, and help resolve an important theoretical debate about the relative roles of deterministic and stochastic processes in community assembly.

Acknowledgements:

We owe enormous thanks to the staff, researchers, students, and interns who have worked for more than three decades to maintain the old field surveys and experiments at Cedar Creek. We are particularly grateful to J. Krueger, T. Mielke, K. Worm, and K. Freund for their help and advice in the 2011-2016 field seasons, and to D. Bahauddin and S. Barrott for assistance with data management and logistics. We thanks E. Borer, C. Lehman, C. Neuhauser, and D. Tilman for helpful comments on earlier drafts of this manuscript. A.T.C was supported through the International Balzan Prize Foundation (awarded to D. Tilman), and by a NSF graduate research fellowship (base award number 00006595). Data collection was supported by the NSF LTER program, including DEB-8114302, DEB-8811884, DEB-9411972, DEB-0080382, DEB-0620652, and DEB-1234162, and by the Cedar Creek Ecosystem Science Reserve and the University of Minnesota. Computing resources were provided by the University of Minnesota Supercomputing Institute.

Figures:

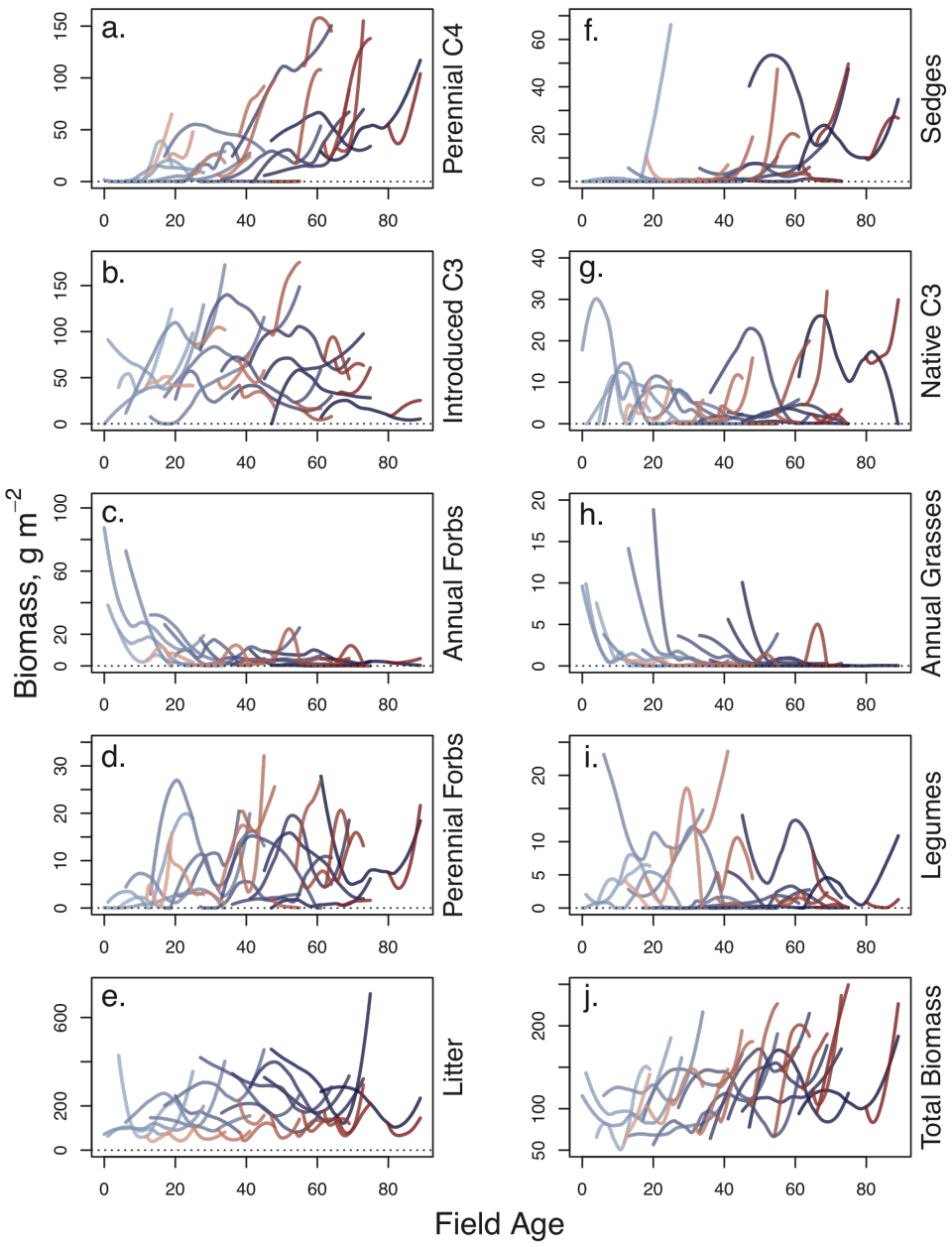

Figure 1: Changes in aboveground biomass as a function of field age (i.e. years since abandonment) for **(a-d,f-i)** various plant species groups, **(e)** leaf litter, and **(j)** total aboveground biomass. Each line shows the successional trajectory for a single field. Red and blue lines show dynamics in burned and unburned transects, respectively. Lighter colours show dynamics for younger fields. Note the difference in scale for the vertical axis among panels.

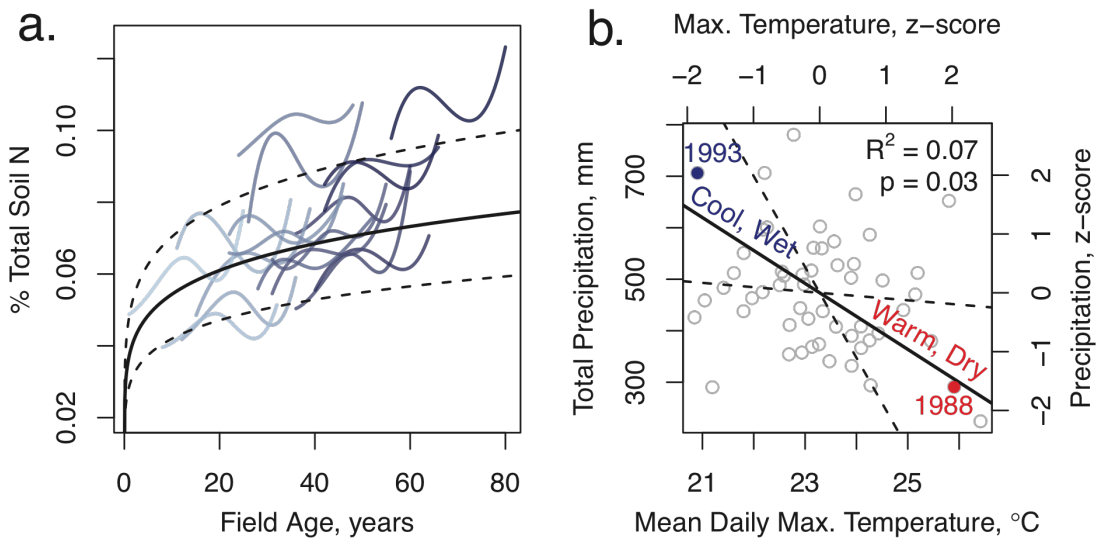

Figure 2: Environmental parameters for the metapopulation model. **(a)** Relationship between % soil total nitrogen and field age. Blue lines show trends for each field based on cubic splines. Black and dashed lines show mean trend ± one standard error, based on the regression described in the main text. **(b)** Relationship between mean daily maximum temperature and total precipitation, measured April-August. Lines, R^2, and p-value show results from ranged major axis regression. Dashed lines show 95% confidence interval for slope. Z-scores show deviations from mean in units of standard deviations. In general, years are either cool and wet, or warm and dry: 1988 “North American Drought” and 1993 “Great Mississippi Flood” are labelled for context.

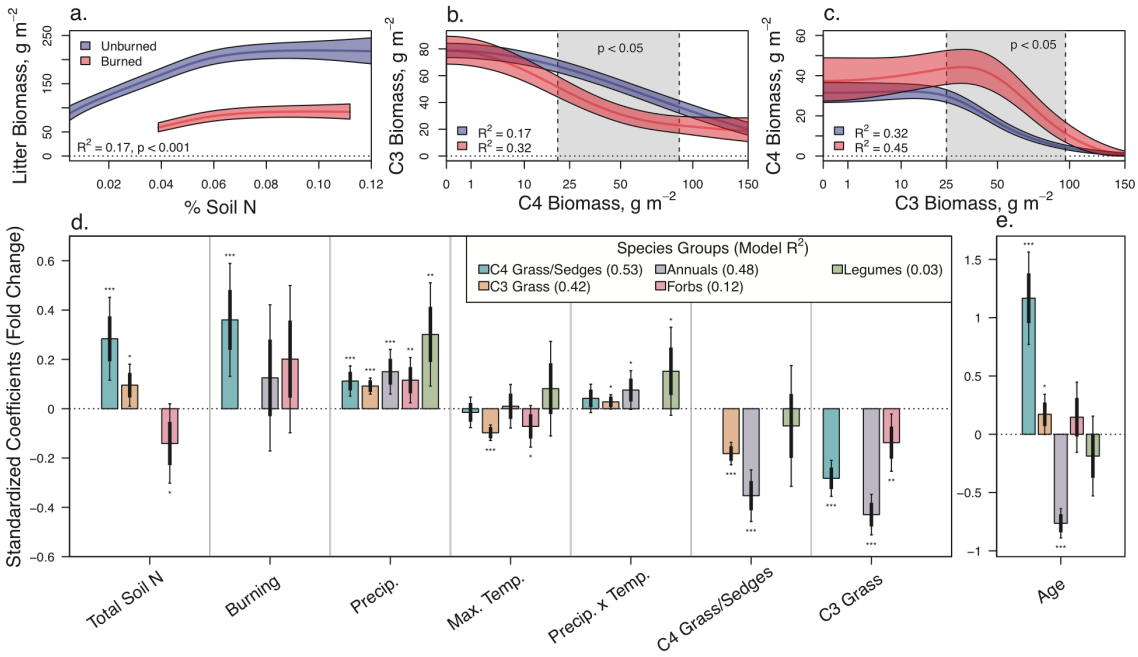

Figure 3: Results for regressions describing plant abundance. **(a)** Effect of burning on leaf litter abundance as a function of soil fertility (p-value tests difference in intercept between treatments). **(b-c)** Relationship between C3 and C4 grass abundance as a function of burning treatment. Grey shaded areas show region with significant differences in abundance between burning treatments. **(d-e)** Standardized coefficients from the abundance regressions. Fold change describes the ratio change in biomass associated with an increase in one standard deviation of each model variable, holding all other variables at their mean value (e.g. fold change of 0 implies no change, 1 implies a doubling). Numbers in parentheses show R^2 for each model. Bars and lines show mean ± one and two standard errors. Asterisks denote significance (***p<0.001; **p<0.01; *p<0.05).

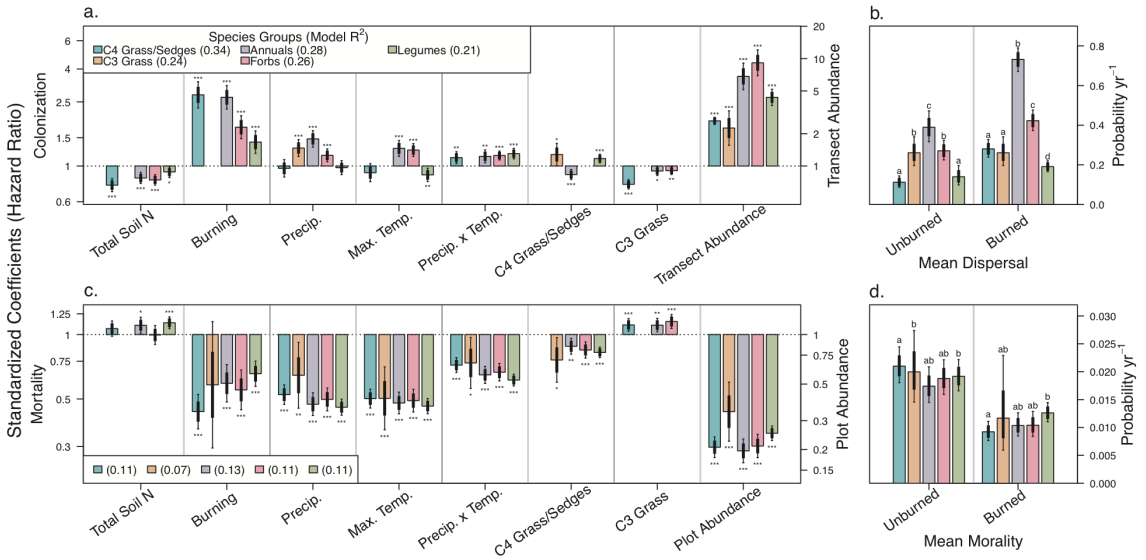

Figure 4: Results for regressions of **(a-b)** colonization and **(c-d)** mortality dynamics. **(a,c)** Hazard ratios show the fractional change in the annual probability of an event associated with an increase in one standard deviation of each model variable, holding all other variables at their mean value (e.g. hazard ratio of 1 implies no change in probability, 1.5 implies a 50% increase in probability). Labels, bars, and lines are as described in the legend for Fig. 3. Numbers in parentheses show R^2 for each model. Note that coefficients for transect and plot abundance are plotted on different vertical axes from the other variables. **(b,d)** Mean annual probability of dispersal and colonization, respectively, in burned and unburned treatments. Letters denote groups that are not significantly different from one another ($p>0.05$).

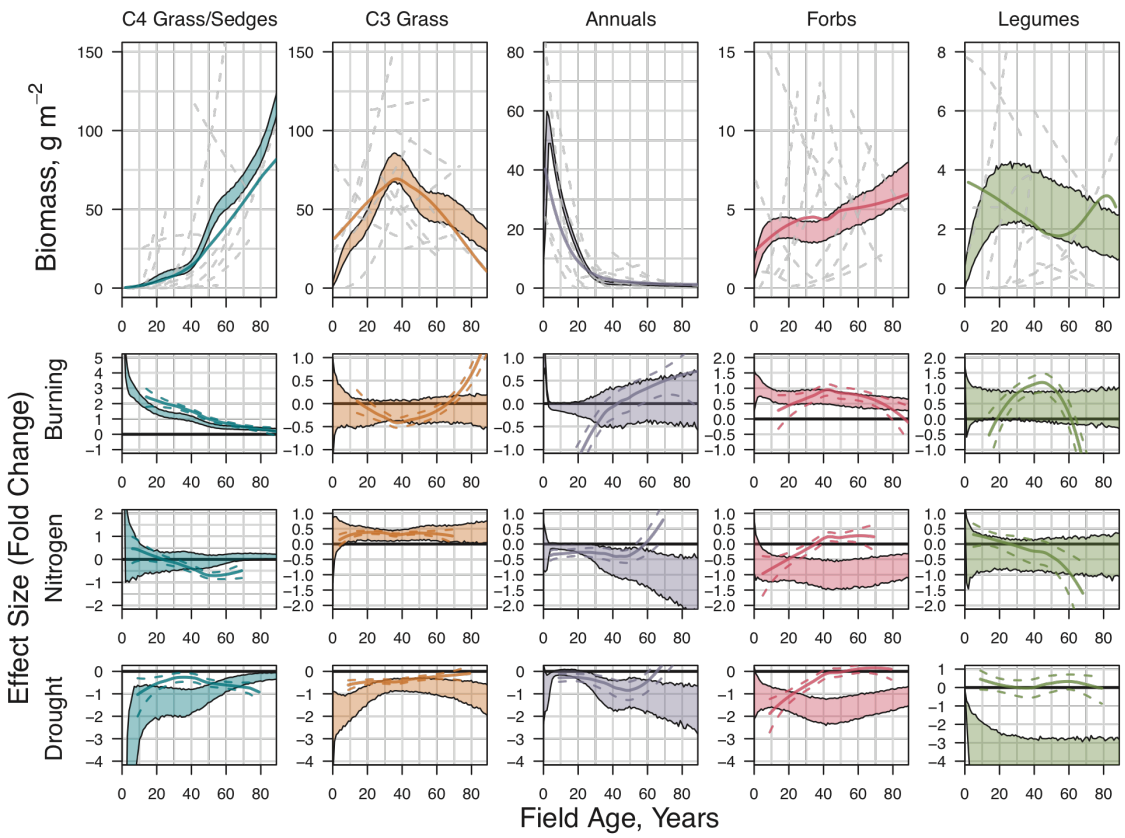

Figure 5: Results from metapopulation simulations. Columns show species groups; rows show scenarios. Shaded regions show mean ± one standard deviation calculated across simulations; solid lines show average trends observed across fields. All results include combined effect of abundance and presence/absence. **(top row)** Reference scenario: aboveground biomass in unburned fields with average nitrogen dynamics and climate conditions. Dashed grey lines show observed trends in each field. **(all other rows)** Difference in biomass among successional trajectories for scenarios, relative to the reference scenario. Dashed lines show mean ± one standard error for observations. **(burning)** Simulations and raw data compare burned vs. unburned fields. **(nitrogen & drought)** Simulations show conditions that are two standard deviations more extreme than the mean; raw data show observations that are more extreme than the mean.

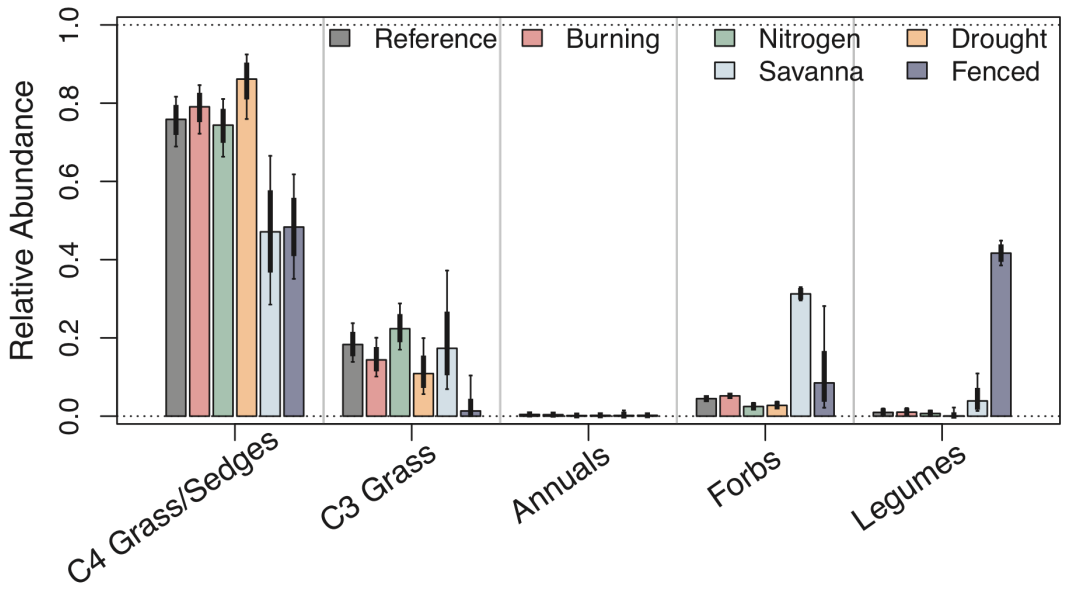

Figure 6: Potential "end states" of successional dynamics. Bars and lines show mean ± one and two standard errors. "Reference", "Burning", "Nitrogen", and "Drought" show results from the four metapopulation simulation scenarios at age 88. "Savanna" shows community composition in a nearby oak savanna that has never been ploughed (from control plots in field "D", described in Tilman 1987). "Fenced" shows community composition in plots from a fenced biodiversity experiment at Cedar Creek, which were seeded with 32 prairie plant species in 1994, and meant to approximate mature prairie (from control plots described in Farrior *et al.* 2013). Note that legume abundance tends to be higher in fenced plots, likely due to selective herbivory (Ritchie & Tilman 1995).

Chapter Two:

Harnessing Uncertainty to Approximate Mechanistic Models

Abstract:

Because Lotka-Volterra competition models posit no specific competitive mechanisms, they are exceedingly general, and can theoretically approximate any underlying mechanism of competition near equilibrium. In practice, however, these models rarely generate accurate predictions in diverse communities. We propose that this contrast between theory and practice may be caused by how uncertainty propagates through Lotka-Volterra systems. In approximating mechanistic relationships with Lotka-Volterra models, associations among parameters are lost, and small variation can correspond to large and unrealistic changes in predictions. We demonstrate that constraining Lotka-Volterra models using correlations among parameters expected from hypothesized underlying mechanisms can reintroduce some of the underlying structure imposed by those mechanisms, thereby improving model predictions by both reducing bias and increasing precision. Our results suggest that this hybrid approach may combine some of the generality of phenomenological models with the broader applicability and meaningful interpretability of mechanistic approaches. These methods could be useful in poorly understood systems for identifying important coexistence mechanisms, or for making more accurate predictions.

Introduction:

Of all of the ecological models used to describe species dynamics, none is so ubiquitous as Lotka-Volterra competition (Lotka 1932; MacArthur & Levins 1967; Wangersky 1978). This model approximates interactions among species as a series of simple, linear functions, describing the per-capita effect of each species on other species growth rates. Lotka-Volterra models are therefore considered the simplest possible abstraction of competition (MacArthur 1970; Tilman 1982). Consequently, Lotka-Volterra competition is typically one of the first models of interspecific interactions taught to students in ecology classes, and underpins an enormous span of ecological

theory, ranging from the criteria for coexistence and competitive displacement (MacArthur & Levins 1967; Chesson 1990, 2000), to relationships among community diversity, productivity, and stability (May 1973; Lehman & Tilman 2000; Loreau 2004).

Besides its relatively simple nature, Lotka-Volterra models are popular because they make no specific assumptions about the mechanisms underlying competitive interactions – i.e. they are entirely "phenomenological." Models that include specific mechanisms necessarily restrict the kinds of interactions that can take place among species. Because Lotka-Volterra models include no such restrictions, they can be parameterized in ways that approximate any combination of underlying mechanisms, at least locally around equilibrium (MacArthur 1970). This can be quite valuable, as it means that Lotka-Volterra models can be applied to a broad array of systems, as opposed to more mechanistic approaches for which the "correct" underlying mechanism may need to be identified in order to make accurate predictions (but see Schaffer 1981). Indeed, there are many classic examples of ecological systems in which dynamics are consistent with the qualitative expectations of Lotka-Volterra models, including aquatic microbial communities (Gause 1934), flour beetles (Park 1936), and warblers (MacArthur 1958), and even some examples where quantitative estimates from these models accurately predict abundances in multi-species communities (Vandermeer 1969; Carrara *et al.* 2015).

Despite this broad applicability, there are substantially more examples of cases where Lotka-Volterra models have failed to accurately predict outcomes of competition. This frequently manifests as a phenomenon known as "non-additivity" or "higher order interactions," in which the effects of two species on one another's growth rates vary depending on the presence of a third species, or of some other factor such as environmental variation (Wilbur 1972; Roxburgh & Wilson 2000; Dormann & Roxburgh 2005; Weigelt *et al.* 2007; Michalet *et al.* 2015). While predictions can occasionally be improved by augmenting models to include parameters that describe these higher-order interactions (Gause 1934; Wilbur 1972; Weigelt *et al.* 2007), predictive power generally remains low, and the changes in model form make them more difficult to generalize to other systems.

The poor performance of Lotka-Volterra models in many real-world ecological systems is hardly surprising. A hallmark of ecological data is that it is highly uncertain – because of both observation error, which results from imperfect measurement of a system, and process noise, which describes natural variability in the characteristics of the system. This uncertainty disproportionately influences Lotka-Volterra models, particularly when they include many species (Dormann 2008; Palamara *et al.* 2016). Lotka-Volterra models abstract systems into a series of linear interactions regardless of the underlying mechanism. For coexistence to be stable in the Lotka-Volterra framework, the Jacobian matrix of the system must have as many negative eigenvalues as there are coexisting species (when evaluated at equilibrium, across all species with positive equilibrium abundance). This requires that the matrix summarizing all pairwise interactions among coexisting species must be of full rank (i.e. all rows and columns must be linearly independent) (Levin 1970; Chesson 1990; Haygood 2002). In other words, coexistence in Lotka-Volterra models depends primarily on the degree of linear independence among the parameters that summarize interactions among species.

However, this method of summarizing and abstracting information is highly vulnerable to uncertainty: even very small additions of error or noise can lead to all interactions "appearing" as though they are linearly independent, and can distorts the original model in ways that are difficult to predict or reverse (Anderson, Guionnet & Zeitouni 2010). This problem is nicely demonstrated in a classical result from May (1973), which shows that large communities with randomly chosen interaction parameters (analogous to noise in the measurement of many parameters) are increasingly unlikely to coexist stably. Similarly, models of communities with random fluctuations in their interaction parameters will drift away from the mechanistic constraints that allow them to coexist, leading to predictions of instability even if the actual system is, in fact, stable.

In many cases, observation error can be mitigated through proper replication and experimental controls (but see *Some notes on observation error* in the Discussion). In contrast, process noise results from variation that is inherent to the system itself (e.g. within-species trait variability, environmental heterogeneity), and cannot be so easily

mitigated. However, because process noise actually influences the dynamics of systems, it may be possible to use this variation to learn about the system's dynamic structure. This is, after all, the primary insight behind regression analysis: if variability is shared among several components of a system, this may indicate that they are meaningfully related. Importantly, because individual mechanisms can only recreate a subset of the types of interactions that can be expressed in Lotka-Volterra models, this suggests that the mechanisms that actually underlie coexistence in a system should leave behind a specific "signature" in the parameters of the Lotka-Volterra model (Tilman 1982). For example, in a model of competition for a single limiting resource, variation in the actual competitive ability of the "best" competitor will alter its interactions with all other species in the community in comparable ways. Conceptually similar, but more complex restrictions have been identified for many classes of mechanistic models that have been converted into Lotka-Volterra form (May & MacArthur 1972, [CSL STYLE ERROR: reference with no printed form.]; Schoener 1974; Tilman 1982).

One of the simplest ways to identify this mechanism-driven correspondence among parameters is through covariance. Covariance measures the degree of correspondence between two variables (correlation is simply covariance standardized by univariate variance). As we will show, the covariance among interaction parameters in a Lotka-Volterra system can be derived directly from a hypothesized underlying mechanistic model, and doing so often requires relatively little information about the specific parameter values and functional form of this mechanistic model. Once calculated, this covariance can be used to constrain the Lotka-Volterra system. Though covariance alone will not fully capture the true dynamics of the underlying mechanistic model, it may be that it can reintroduce some aspects of the structure that these mechanisms impose on system dynamics. Modeling covariance could therefore be useful for improving predictions, and identifying signatures left behind by influential coexistence mechanisms.

In this manuscript, we demonstrate the potential utility of harnessing covariance to detect and integrate components of mechanistic models into the phenomenological Lotka-Volterra framework. As a worked example, we use a model developed by

MacArthur (1970), which relates a mechanistic resource competition model to the classical Lotka-Volterra framework. After *(1)* introducing this model, we will use it to *(2)* demonstrate how process noise in the mechanistic parameters propagates through the system and changes model predictions, *(3)* derive the covariance relationship among phenomenological parameters that results from this process noise, *(4)* show how this information can reduce distortion of the model, and *(5)* explore how this theoretical insight can be applied in the analyses of empirical data to identify underlying mechanisms and improve model predictions.

Model and Results:

1. MacArthur's resource competition model

MacArthur's 1970 resource competition model relates a mechanism-based model of interspecific competition for perfectly substitutable resources to the classical Lotka-Volterra competition framework (MacArthur 1970; Tilman 1982). MacArthur's simplifying assumption was to propose that resource dynamics occurred much faster than consumer dynamics, which allowed resource concentrations to be estimated as a simple function of consumer abundance. Because of this simplification, the model can re-written in a way that is mathematically identical to the classical Lotka-Volterra competition equations, despite the fact that it includes parameters that can be interpreted mechanistically (though see caveats in *Potential limitations* in the discussion). In this model, interspecific interaction terms of the Lotka-Volterra competition equations (a_{ij}, which describe the per-capita effect of species j on the growth rate of species i) can be directly related to a function of the mechanistic parameters (c_{il}, which describe the ability of species i to acquire resource l). Note that we use the notation of Chesson (1990).

This model has been extensively developed in many subsequent studies (e.g. Schoener 1974; Chesson 1990, 2000; Haygood 2002), and has been found to be particularly tractable for a number of reasons. First, any combination of n species and l_{max} limiting resources for which a stable, non-trivial equilibrium exists (i.e. not all species are at zero abundance), the model is globally stable, meaning that is approaches a single equilibrium from any starting point. Second, competition is symmetrical between species

(i.e. $a_{ij} = a_{ji}$), which provides a specific structure to competitive interactions and reduces the number of parameters. Lastly, and most importantly for our purposes, the model provides a simple link between "mechanistic" parameters describing species consumption rates (c_{il}), and "phenomenological" parameters describing the outcomes of competitive interactions (i.e. a_{ij}). These parameters are often grouped into matrices $\boldsymbol{c}$ and $\boldsymbol{a}$, respectively.

To facilitate the process of tracking error propagation, we further simplify MacArthur's original model (full derivation of our model, and a more detailed discussion of how the parameters relate to other ways of expressing the Lotka-Volterra system, are available in Appendix A.I in the supplement.) In our framework, $\boldsymbol{c}$ and $\boldsymbol{a}$ are related following

$$a_{ij} = \sum_{l=1}^{l_{max}} c_{il} c_{jl} \qquad \text{Eq. (1)}$$

Dynamics in the abundance of species i, X_i, depend only on the species growth rate, b_i, the maximum amount of resource taken up in the absence of other competitors, k_i, and competitive interactions with other species, following the form

$$\frac{1}{X_i}\frac{dX_i}{dt} = b_i\left(k_i - \sum_{j=1}^{n} a_{ij} X_i\right) \qquad \text{Eq. (2)}$$

Eq. (2) is mathematically identical to the classical Lotka-Volterra competition equations, though with mechanistic constraints on the values of $\boldsymbol{a}$. Note, however, that we use a different parameterization in Eq. (2) than in presented in many ecological textbooks (e.g. Gotelli 2008). Unlike these "classical" forms, carrying capacities and interaction coefficients in our model are not standardized by species self-inhibition rates. Thus, the "classical" carrying capacity K_i (i.e. species abundance in the absence of competitors) is equal to k_i/a_{ii} in our model's parameterization, and self-inhibition in our model a_{ii} (i.e. the effect of species i on itself) is not necessarily equal to one (unlike the "classical" α terms, for which $\alpha_{ij} = a_{ij}/a_{ii}$) (Chesson 2000).

For a stable equilibrium to occur in this system, there must be at least as many limiting resources as there are coexisting species (i.e. $l_{max} \geq n$), and consumption vectors of each species, composed of the columns of $\boldsymbol{c}$, must be linearly independent (MacArthur 1970; Chesson 1990). Because $\boldsymbol{a}$ is composed of pairwise products of $\boldsymbol{c}$, satisfying these conditions requires that all rows and columns of $\boldsymbol{a}$ are linearly independent (i.e. that

matrix $\boldsymbol{a}$ is of "full rank") (Chesson 1990). Note, however, that not all communities that meet these criteria will coexist stably (Chesson 1990; Haygood 2002). A full treatment of stability criteria, which also depend on k_i and b_i, is available in Appendix A.II in the supplement.

2. Effects of process noise on model predictions

We define process noise as meaningful variation in the mechanistic consumption parameters of $\boldsymbol{c}$. By "meaningful," we mean that this variation actually alters the consumption vectors themselves, rather than merely changing our ability to accurately measure these parameters (i.e. in contrast to observation error). Meaningful variation might be caused by genetic differences, trait plasticity, spatial heterogeneity, or any other such process that drives within-species trait differences. Process noise therefore can lead to changes in species dynamics, equilibrium population abundances, and even persistence.

Process noise also leads to variation in $\boldsymbol{a}$, because these terms are related to the elements of $\boldsymbol{c}$ following Eq. (1). As an example, consider a system where process noise is normally and independently distributed around each of the elements of $\boldsymbol{c}$, with standard deviation σ_c. Based on the mechanistic relationship between $\boldsymbol{c}$ and $\boldsymbol{a}$ in Eq. (1), process noise in $\boldsymbol{c}$ can be analytically related to variation in the terms of $\boldsymbol{a}$ as

$$\sigma_{a_{ii}}^2 = 2{\sigma_c}^2\left(l_{max}{\sigma_c}^2 + 2a_{ii}\right) \qquad \text{Eq. (3a)}$$

$$\sigma_{a_{ij}}^2 = {\sigma_c}^2\left(l_{max}{\sigma_c}^2 + a_{ii} + a_{jj}\right) \qquad \text{Eq. (3b)}$$

where σ_{aii} and σ_{aij} are the standard deviation in the terms of $\boldsymbol{a}$ resulting from σ_c (derivations in Appendix B.I in the supplement). These terms increase roughly as a linear function of σ_c (Fig. 1a, see linear approximations in Appendix B.I), and are of a similar magnitude as σ_c regardless of community size (Fig 1b).

For simplicity, let us suppose that we can perfectly measure all of the parameters in this system, and that the terms k_i and b_i are not subject to any kind of process noise. In this case, variability in the outcomes of competitive interactions can be entirely attributed to σ_c and its effects on the realized values of $\boldsymbol{c}$ and $\boldsymbol{a}$. As an example, consider a case of two species competing for two limiting resources which are able to coexist in the absence

of process noise (i.e. $\sigma_c = 0$). If we increase process noise in this system, the increasing variability in $\boldsymbol{c}$ and $\boldsymbol{a}$ would lead to increasing probability of competitive exclusion between the two species (Fig 2a). For example, process noise resulting from increasing spatial heterogeneity could lead to competitive displacement in a greater fraction of sites that the species jointly occupy.

Now, suppose that we attempt to predict the outcome of competition by directly measuring the components of $\boldsymbol{a}$. Most commonly, this is accomplished by measuring a_{ii} and a_{jj} under circumstances where only one species is present, and a_{ij} and a_{ji} under circumstances where both species are jointly present. In this case, we would find something curious: even if we perfectly measure the components of $\boldsymbol{a}$ in the presence of process noise, we will overestimate the frequency of competitive exclusion (Fig 2b). This is because we fail to take into account the correlation among the components of $\boldsymbol{a}$. However, note that if we account for covariance among these terms (for example, by measuring the terms at the same time in the same system), we are able to much more closely match the outcome expected from the mechanistic model.

3. Model covariance relationship

The correlations relating elements of $\boldsymbol{a}$ impose limitations on its structure, thereby restricting the very general form of Lotka-Volterra competition to behave somewhat more like the specific mechanisms posited in the MacArthur model. As the number of competing species grows, these relationships become increasingly complex and influential. In contrast to the example in Fig. 2, higher-dimensional correlations may lead to increases or decreases in the estimated number of coexisting species. It is therefore helpful to derive the expected covariance relationship among all terms in $\boldsymbol{a}$ based on their mechanistic relationships in the MacArthur model. This reveals six classes of elements, including the two variance terms in Eqs. (3a-b)

$$\mathrm{cov}(a_{ij}, a_{ji}) = {\sigma_c}^2\left(l_{max}{\sigma_c}^2 + a_{ii} + a_{jj}\right) \quad \text{Eq. (3c)}$$

$$\mathrm{cov}(a_{ij}, a_{ii}) = 2{\sigma_c}^2 a_{ij} \quad \text{Eq. (3d)}$$

$$\mathrm{cov}(a_{ij}, a_{ik}) = {\sigma_c}^2 a_{jk} \quad \text{Eq. (3e)}$$

$$\mathrm{cov}(a_{ij}, a_{km}) = 0 \quad \text{Eq. (3f)}$$

where i, j, k, and m represent four distinct species. Note that because of symmetry in $\boldsymbol{a}$, these classes include several kinds of associations (e.g. $\text{cov}(a_{ij}, a_{ii}) = \text{cov}(a_{ij}, a_{jj})$). Full derivations and details are available in Appendix B.I in the supplement.

4. Testing the covariance relationship

We can demonstrate how the covariance relationships capture some aspects of the mechanistic model by simulating three types of models with added process noise: *(i)* the mechanistic model based on Eqs. (1-2); *(ii)* an uncorrelated model, in which variance in $\boldsymbol{a}$ is calculated following Eqs. (3a-b), but covariance is ignored; and *(iii)* a covariance model, which accounts for all variance and covariance relationships among components of $\boldsymbol{a}$ following Eqs. (3a-f). Thus, the mechanistic model represents the "true" process, the uncorrelated model represents a sampling design that ignores correlations among species interaction coefficients, and the covariance model demonstrates potential improvements in predictions resulting from properly accounting for the mechanistic model's effects on the correlation structure of $\boldsymbol{a}$.

To demonstrate the effects of community diversity and noise on these three models, we simulated each of them across five sizes of communities (n = 2, 4, 6, 8, 10), and four levels of process noise (σ_c = 0.01, 0.05, 0.1, 0.15). For each of these scenarios, we identified stable equilibria, and compared model predictions for species richness, species abundance, and for the elements of $\boldsymbol{a}$. For all scenarios, we set $l_{max} = n/2$, and ran 20,000 iterations for each model. Detailed methods for these three models are available in Appendix C in the supplement.

We find that for predictions of community richness, the covariance model matches the mechanistic model much more closely than does the uncorrelated model, both in terms of increased precision (i.e. less variance around the mean prediction) and decreased bias (i.e. the predictions are more centered around the true values) (Fig. 3). Correlation between estimates of $\boldsymbol{a}$ from the three models remains relatively high across all models and scenarios. For large community sizes and large process noise (e.g. n = 10 and σ_c = 0.15), the uncorrelated model also tends to over-predict coexistence, while the

covariance model does not. On average, the covariance model also provides better predictions of abundance, but the difference is relatively small.

5. Applying theoretical results

The results from Fig. 3 demonstrate that we can more accurately predict outcomes of the MacArthur model by incorporating the mechanistic signature of covariance imposed on the interaction matrix $\boldsymbol{a}$ than we can by measuring the components of $\boldsymbol{a}$ independently. Note that this procedure assumes that we already know the "correct" expected values of all of the terms in $\boldsymbol{a}$, and the magnitude of σ_c. However, under these circumstances we would often have enough information to reconstruct $\boldsymbol{c}$, at which point it would be more efficient to simply make predictions based on the underlying mechanistic model. To make our results somewhat more useful for real-world applications, we demonstrate a series of methods in the section blow that could be used to parameterize the covariance model based on empirical observations of species communities, even when the true values of $\boldsymbol{a}$ and $\boldsymbol{c}$ are not known.

5.1 Detecting model dimensionality

The first challenge is to identify the number of limiting resources in the system (but see Appendix B.I for approximations that can sometimes avoid this necessity). In theory, the number of nonzero eigenvalues of $\boldsymbol{a}$ should indicate its rank. In practice this is less straightforward because variation in the elements of $\boldsymbol{a}$ in the uncorrelated model will artificially increase the dimensionality of $\boldsymbol{a}$ to full rank regardless of l_{max} (Anderson *et al.* 2010; Bates & Maechler 2016). However, if process noise is small, then it is possible that these added dimensions will only be weakly present in $\boldsymbol{a}$, in which case they might be identified as corresponding to particularly small eigenvalues, or as components of $\boldsymbol{a}$ that can be removed without worsening model predictions.

For example, in a simulated system with four species, but only two limiting resources, the "true" dimensionality of $\boldsymbol{a}$ should not exceed two, and the third and fourth eigenvalues should be equal to zero (n.b. these eigenvalues are calculated from the

interaction matrix as a metric of dimensionality, not from the Jacobian as a metric of stability). Encouragingly, we find that these eigenvalues from the uncorrelated model are disproportionately small, and centered around zero (Fig. 4a). Furthermore, if we transform this matrix to set the value of these eigenvalues to zero, which effectively removes some of the effects of the uncorrelated noise from the matrix, this leads to improved predictions of richness and abundance, but has very little effect on the actual values of the elements of $\boldsymbol{a}$ (Fig. 4b-d). Detailed methods are described in Appendix D.I in the supplement.

In a system with unknown dimensionality, similar results would be a good indication that the information associated with these eigenvalues was not mechanistically meaningful, and could be useful for detecting the true dimensionality of $\boldsymbol{a}$. Note, however, that this technique is probably not a good method for making predictions directly from the uncorrelated model. First, we find that the improvement in prediction power is smaller than that achieved by the covariance model. Second, as systems grow larger, this correction technique becomes more difficult to apply because the transformation begins to introduce imaginary parts into $\boldsymbol{a}$, and because rounding errors tend to make it impossible to fully remove many eigenvalues simultaneously.

5.2 Estimating covariance model parameters

Once the true dimensionality of the system is known, the second challenge is to determine whether the parameters of the covariance model can be properly estimated by fitting the model to observed data. If we parameterize the covariance model with values centered around the "true" expected values of $\boldsymbol{a}$ from the mechanistic model, we find that these correspond closely to the average predictions from the mechanistic model (Fig. 5a-f). Similarly, as outlined above, differences in the distribution of eigenvalues resulting from these parameterizations of $\boldsymbol{a}$ successfully identify systems with different numbers of limiting resources (Fig. 5g). Moreover, if we calculate the likelihood of observed outcomes from the mechanistic model (or calculate the likelihood of empirically observed abundances of species from communities that follow the dynamics of the

mechanistic model), we find that likelihood increases as the estimated $\boldsymbol{a}$ terms approach the mechanistic values (Fig. 5g-h).

Jointly, these results show that by optimizing parameters in the covariance model such that they maximize the likelihood of observed data given the covariance model, estimates should converge on the true values of $\boldsymbol{a}$. These analyses therefore provide a proof of concept that regression or optimization tools that are able to estimate components of a covariance matrix (e.g. GLS, Bayesian hierarchical modeling) could be used to empirically parameterize the covariance model based on observed data. Full methods for this procedure are described in Appendix D.II in the supplement.

Discussion:

Our results demonstrate two primary points. First, we show that predictions from Lotka-Volterra models that do not account for mechanistic associations among model parameters will often be inaccurate and biased, particularly for communities with many competing species. Second, we find that underlying mechanistic relationships among the interaction parameters of a Lotka-Volterra system can be successfully approximated using analytically derived covariance, which helps improve predictions and ameliorate bias. These results therefore show that naively parameterizing Lotka-Volterra models from field data will often lead to models with poor predictive power. However, our findings also suggest that at least in some systems, it may be relatively straightforward to parameterize a semi-mechanistic "hybrid model," which uses covariance to incorporate the rough skeleton of hypothesized coexistence mechanisms, but does not require as much detailed information as would a fully mechanistic approach.

Note that poor performance of the uncorrelated model is not contingent on uncertainty arising from process noise. Any variation that is not constrained by underlying mechanistic relationships will deform the interaction matrix $\boldsymbol{a}$, resulting in increased dimensionality (Anderson *et al.* 2010; Bates & Maechler 2016), and therefore poorer predictions (Dormann 2008). Thus, even if uncertainty arises from observation error rather than process noise, methods that do not constrain $\boldsymbol{a}$ to retain its mechanistic

structure will likely fail when making predictions for communities of more than a few species.

Though models that incorporate true underlying mechanisms will likely generate predictions that are more accurate and generalizable than those from covariance approximations, there are nevertheless some advantages to the hybrid approach. Note that Eqs. (3a-f) do not require knowledge of species consumption rates, nor the identity of the resources for which species compete. Likewise, the number of limiting resources can often be estimated directly from $\boldsymbol{a}$ (Fig. 4, and Appendix D.II). Thus, with relatively limited mechanistic information, it may nevertheless be possible to derive and parameterize the covariance model. For example, in any system where we suspect that species interactions may be guided by resource competition as defined in the MacArthur model, the covariance model described here could be used to approximate system dynamics, even if the precise number and identity of limiting resources were not known.

The potentially broad applicability of covariance models raises the question of whether covariance among elements in $\boldsymbol{a}$ could be measured entirely empirically, rather than deriving it from a hypothesized mechanistic model. We suspect not. There are n^4 terms describing covariance among the elements of $\boldsymbol{a}$, and measuring many of these components (e.g. Eqs. (3e-f)) would likely require replicated observations of all possible three-way combinations of species, which increase as $n!/(6(n-3)!)$. While there are some examples of studies that realize this level of replication (Wilbur 1972; Miller 1994; Weigelt *et al.* 2007), these are limited to communities of relatively few species. Furthermore, these sorts of experiments are generally only feasible for study species that are small and fast-growing. One possible approach to reduce replication requirements might be to estimate interaction coefficients from time series recorded in multi-species communities, but confounding factors such as joint responses to environmental variation would need to be controlled for (Deyle *et al.* 2016). Thus, while empirically calibrated covariance might serve as a preliminary test to narrow down a large list of potential coexistence mechanisms, we think it is very unlikely that such a method could be used to improve model predictions without first specifying a mechanistic "backbone."

Interestingly, covariance could potentially explain non-additivity observed in some other studies (Wilbur 1972; Miller 1994; Dormann & Roxburgh 2005). For example, consider the terms a_{ij} and a_{ik}, describing the effect on species i of species j and k, respectively. Because $\text{var}(a_{ij} + a_{ik}) = \text{var}(a_{ij}) + \text{var}(a_{ik}) + \text{cov}(a_{ij}, a_{ik})$, individual observations of the joint effect of these species will either be weaker or more extreme than would be expected from two-way interactions. Unless replication of three-way interactions is sufficiently large, covariance would manifest as a change in the strength of competitive interactions depending on the presence of a third species. Thus, studies that have identified non-additive competitive interactions in the past may, in fact, constitute further evidence for the important role of covariance in Lotka-Volterra models.

Potential limitations

Though simple and relatively easy to interpret, an important caveat for the implementation of MacArthur's resource competition that we utilize here is that it is in many ways only "semi-mechanistic". An important underlying assumption within this model is that resources are "perfectly substitutable" in this model – i.e. any species can theoretically persist on a sufficient quantity of any single resource. Furthermore, the combination of traits that we generate for species (and that are usually utilized for this model) assumes that all species forage for multiple resource simultaneously. Lastly, the model assumes strictly linear relationships among resource requirements (i.e. the effect on individual species of adding one type of resource to the system is always a fixed fraction of the effect of adding another type of resource, regardless of the available concentration of either resource). In reality, these assumptions are relatively unrealistic, as few resources are actually perfectly substitutable, competition among substitutable resources tends to lead towards "switching" behavior (i.e. species specializing on harvesting a single type of resource), and species responses to most resources tend to saturate at higher availabilities, leading to nonlinear responses (Tilman 1982). Because of these caveats, we would expect evolution to drive species in our model to become specialists on individual resources. Thus, though helpful as a worked example, we would

not expect this model to accurately predict community dynamics in most real-world systems.

A potential problem with the covariance approach that we use here is that in some cases, we may find that mechanistic systems cannot be well-approximated merely by incorporating covariance. Because we characterize the components of $\boldsymbol{a}$ entirely by their mean, variance and covariance, we technically are assuming that these terms are drawn from a shared multivariate normal distribution. However, the terms for $\boldsymbol{a}$ in the MacArthur model are technically the result of product distributions (Grimmett & Stirzaker 2001). While product distributions – and other non-normal distributions – can sometimes be approximated by normal distributions, they can also be very difficult to work with because they are not always symmetrical, and can include "higher-order" moments (e.g. skew or kurtosis). Though we were able to generate accurate predictions for the MacArthur model without including these higher-order moments (i.e. predictions from our covariance model converged around those for the mechanistic model), more complex distributions may be needed to accurately characterize other types of mechanistic models. This is a well-known problem in analytical models of dispersal, where "spatial moments" (akin to variance, covariance, etc., but describing correlations across space) are used to approximate spatial dynamics (Bolker & Pacala 1999). Though simple spatial moment approximations work well for some systems, others require such complex approximations that it can be simpler to identify and parameterize the actual underlying mechanistic model (Murrell, Dieckmann & Law 2004).

Another potential problem with our approach is model identifiability. Because covariance includes relatively little information about mechanistic structure, there may be multiple models that generate similar covariance signatures. For example, both neutral theory and Connell's intermediate disturbance hypothesis maintain diversity through transient dynamics (Wright 2002), and the covariance structures of Lotka-Volterra abstractions of these mechanisms may consequently be difficult to distinguish. Thus, while identifying covariance among Lotka-Volterra terms may be useful for improving predictions given a hypothesized underlying mechanism, or for narrowing down the range of mechanistic models that could potentially describe observed dynamics,

covariance is likely not the best approach for definitively identifying which of many similar mechanisms is most likely to have generated observed patterns.

Nevertheless, this limitation could also be a useful property of our method. For example, mechanistic models of disease propagation (May & Anderson 1987) and plant metapopulation dynamics (Tilman 1994) can be developed with identical relationships among model parameters, despite obvious differences between the systems. Lotka-Volterra abstractions for these models would therefore also have identical covariance. In less obvious cases, such correspondences might be helpful for identifying seemingly disparate, but mechanistically related, classes of models.

Broader implications

Though we use one version of the MacArthur model as an example, the underlying methods we present here could be applied to other mechanistic models. Even where analytical derivation relating process noise to variation in phenomenologically observable parameters is not practical, the same results can be achieved by empirically calculating covariance from simulations of the posited underlying mechanistic model. Importantly, these relationships should be derived independently for each model mechanism that utilizers of this approach wish to test. We therefore caution that the specific analytical forms that we derive here should in no way be thought of as "general" across all types of ecological systems.

A remaining, unexplained result is our prediction that process noise leads to changes in species abundances even in the mechanistic model (Fig. 3). We also find moderate among-simulation variation in predictions from the mechanistic model, roughly matching that from the covariance model (Fig. S2 in the supplement). While this prediction is potentially consistent with strong effects of stochasticity (Palamara *et al.* 2016), or of patchy and temporally variable species distributions (Clark, Rykken & Farrell 2011; Li *et al.* 2016), it contrasts with results which show consistent species abundances across replicates in competitive systems (Wedin & Tilman 1993; Harpole & Tilman 2005; Dybzinski & Tilman 2007).

Potentially, variability caused by real-world process noise is constrained by physiological tradeoffs among species traits. Tradeoffs induce stable coexistence in mechanistic models (Tilman 2011), and by reducing the kinds of changes to traits that are physically possible, tradeoffs might prevent process noise from causing overly large changes in species abundances. In our model, we impose tradeoffs among species consumption vectors, but do not constrain process noise to adhere to this surface, as this would have produced covariance among the elements of $\boldsymbol{c}$, making the system analytically intractable (see Appendix C in the supplement). Note, however, that because process noise is relatively small, species traits always fall closely around the tradeoff surface. This is equivalent to assuming that the underlying physiological tradeoffs that constrain species traits in our system are higher dimensional than is the ecological trait space that determines coexistence. Nevertheless, more restrictive tradeoffs in process noise could be incorporated into computational estimates of covariance, which might generate more stable estimates of species abundances, and improved predictive ability.

Some notes on observation error

As previously explained, we primarily discuss the effects of process noise rather than observation error because observation error can often be mitigated. Nevertheless, some kinds of variables are by nature *pathological* – that is, they have no mean or variance (or sometimes even higher moments). Regrettably, ratios of normally distributed variables, which commonly arise in Lotka-Volterra systems, are an important example of such a variable (n.b. the Cauchy distribution is a subtype of ratio distribution) (Marsaglia 2006). For example, competition coefficients are typically estimated as $\alpha_{ij}=(K_i - X_i)/X_j$, growth rates are often calculated as $r_i=\ln(X_i(time=\tau)/X_i(time=0))(1/\tau)$, and even the Jacobian that we use to determine stability includes ratios of random variables (Eqs. (SA12a-b)). Because small changes in the denominator or numerator can cause a ratio to jump between zero and infinity, even very large sample sizes may not lead to a "stable" estimate of the variable's distribution: the millionth sample can cause a large increase in the estimated mean, while the millionth and first might cause a large decrease.

There are some strategies for reducing the influence of these ratios. If the mean of the numerator and denominator are of a suitable magnitude relative to their variance, then their ratio may be roughly normally distributed, though the precise conditions for this are not trivial (Marsaglia 2006). Alternatively, by fitting models to dynamic data, it may be possible to estimate some of these parameters directly, rather than as a ratio of empirically measured variables (Carrara *et al.* 2015; Palamara *et al.* 2016). Nevertheless, some variables are necessarily the outcome of ratios. For example, coexistence in the MacArthur model depends on the relative consumption rates and carrying capacities of species, not their absolute magnitude (Chesson 1990; Haygood 2002). It therefore seems unavoidable that Lotka-Volterra-like methods (and likely a great many other models of species interactions) will be especially susceptible to the effects of observation error. We therefore advocate cautious testing for the effects of observation error before attributing uncertainty to process noise.

Resource-based models of competition are less subject to the potential impacts of pathological variables because species competitive abilities can be measured as levels of unconsumed limiting resources in monocultures, and are often directly constrained by tradeoffs (Tilman 1994, 2011). Similarly, carrying capacities can be often be directly measured in monocultures, and net resource acquisition (consumption vectors) can be determined by measurement of the concentrations of limiting nutrients in the biomass of each species in monoculture. Thus, ratios of variables are less likely to be required to parameterize such a model.

Conclusion:

Our findings suggest that a simple hybrid approach that tracks the propagation of uncertainty through ecological systems might be useful for identifying important coexistence mechanisms and predicting species abundances in poorly understood competitive communities. It remains to be seen whether the simple approach that we use based on covariance will be tractable and effective for other types of mechanisms and in real-world ecological systems. However, we hope that the methods and concepts that we introduce here both provide a warning of ways that Lotka-Volterra models can be mis-

calibrated, and will help expand the utility of Lotka-Volterra approaches in diverse systems, and help guide how ecologists use these approaches in the future.

Acknowledgements: We thank E. Borer, G. Fury, F. Isbell, C. Lehman, D. Tilman, and D. Williams for helpful comments on earlier drafts of this manuscript. A.T.C. was supported by a NSF Graduate Research Fellowship, base award number 00006595. Computing time was provided by the University of Minnesota Supercomputing Institute.

Figures:

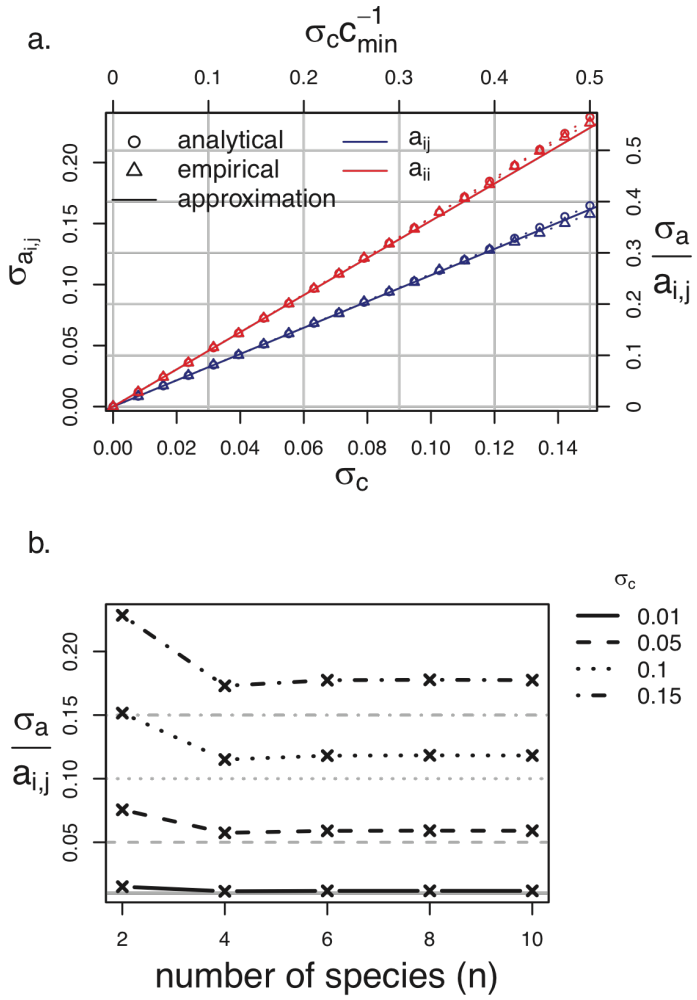

Figure 1: Propagation of process noise in model parameters. **(a)** $\sigma_{ai,j}$ shows standard deviation of the competition coefficients $a_{i,j}$ (blue) and $a_{i,i}$ (red) as a function of σ_c, the standard deviation of noise added to the mechanistic consumption parameters, $\boldsymbol{c}$, in MacArthur's resource model. Circles show analytical expectations for the relationship, triangles show average results from 20,000 simulations of the mechanistic model, and lines show linear approximations, as described in the Appendix B.I in the supplement. Additional axis labels $\sigma_c(c_{min})^{-1}$ and $\sigma_a(a_{i,j})^{-1}$ show the magnitude of these variabilities relative to the minimum $\boldsymbol{c}$ value and mean $a_{i,j}$ value, respectively. **(b)** Mean observed value of $\sigma_{ai,j}$ as a function of community size, n ($l_{max} = n/2$ in all cases). With a slight exception around n=2, $\sigma_{ai,j}$ is approximately equal to σ_c (corresponding values of σ_c shown in grey). Note that the mean value of $\boldsymbol{c}$ is constrained to equal 1 regardless of n, while mean of $a_{i,j}$ increases roughly as $n/2$ (not shown).

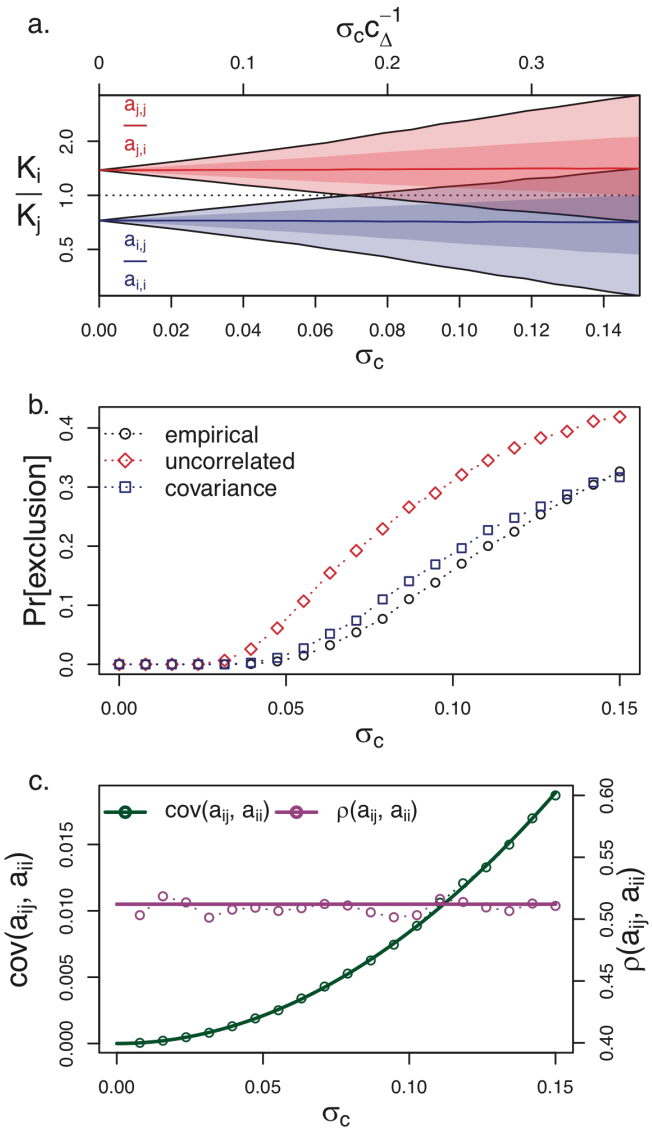

Figure 2: Effects of process noise on coexistence for a system with two species and two limiting resources. **(a)** Species *i* and *j* coexist when the ratio of their carrying capacities K_i/K_j falls between a_{ij}/a_{ii} and a_{jj}/a_{ji} (i.e. between the dark red and dark blue lines – see Fig. 1 in Chesson (1990) for a similar approach). Lines, dark shaded intervals, and light shaded intervals show the mean, standard deviation, and 95% confidence interval, respectively, for these ratios as a function of process noise, σ_c, while dotted line shows fixed value for K_i/K_j. Top axis shows magnitude of noise relative to the mean difference between species consumption rates for the two limiting resources. **(b)** Effect of process noise. Pr[exclusion] shows the probability that one of two competing species will drive its competitor extinct. Empirical results show the average of 20,000 simulations of the mechanistic model, while uncorrelated and covariance predictions are based on the analytical expectation of variance and covariance in $\boldsymbol{a}$, respectively, as described in the main text. Distance from the empirical estimates (black points) demonstrates prediction error. **(c)** Covariance and correlation between a_{ij} and a_{ii} as a function of process noise. Points show empirical results based on the average of 20,000 simulations, while lines show the analytical expectation, as described in the main text.

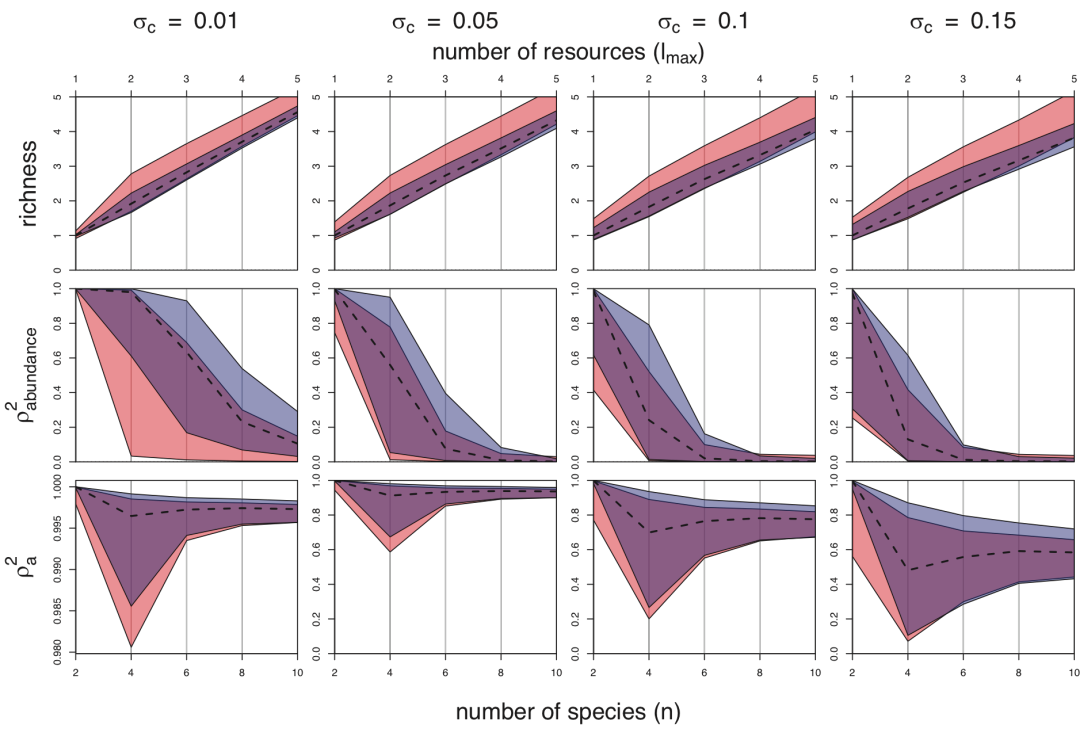

Figure 3: Correspondence between mechanistic model and Lotka-Volterra approximations as a function of process noise, σ_c, for communities of 2 to 10 species (note that $l_{max} = n/2$ in all cases). Intervals show mean ± one standard deviation based on 20,000 simulations. Red intervals show results for uncorrelated model, blue shows covariance model, and purple regions show overlap between the two. Black dashed line shows mean result from the mechanistic model. $\rho^2_{abundance}$ and ρ^2_a show the square of Pearson's correlation coefficient comparing predictions of species abundances or predictions of the interaction matrix $\boldsymbol{a}$, respectively. Note that the vertical axis for ρ^2_a for $\sigma_c = 0.01$ is on a different scale from the other panels.

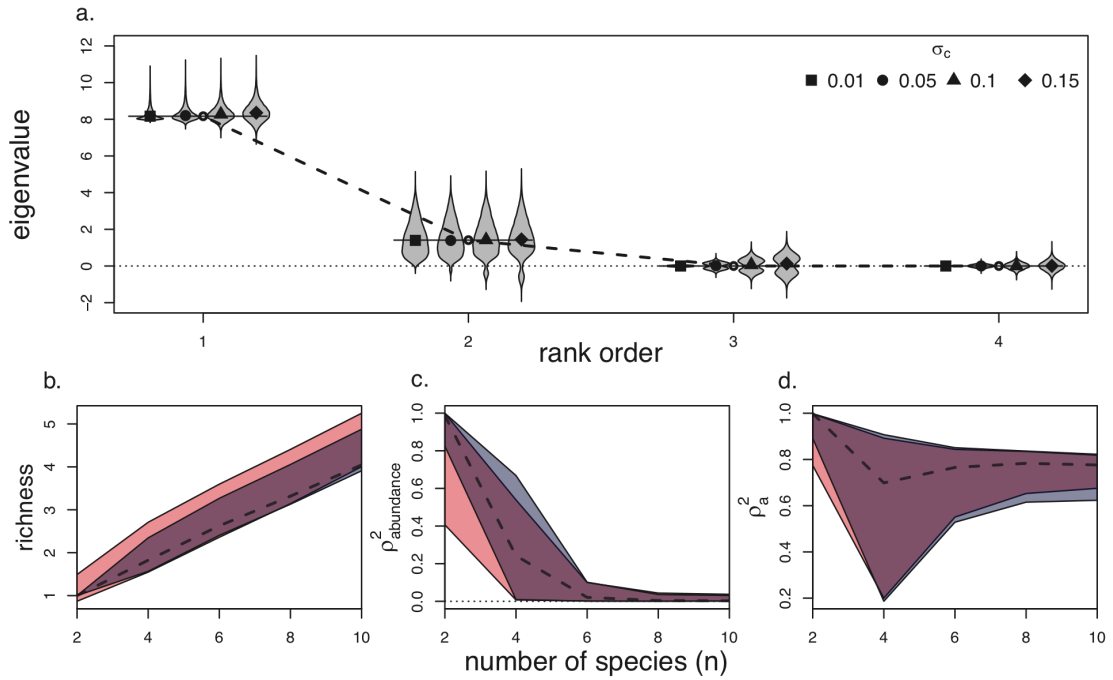

Figure 4: Influences of uncorrelated error on predictions of community stability. **(a)** Real components of the eigenvalues for the interaction matrix $\boldsymbol{a}$ in a system with four species and two limiting resources. Black horizontal lines and dashed segments show mean values for the mechanistic model (n.b. fourth and fifth eigenvalues are always = 0). Width of shaded intervals shows frequency distribution of results for Lotka-Volterra systems with uncorrelated error (i.e. the uncorrelated model). **(b-d)** Correspondence between mechanistic model and Lotka-Volterra approximations for 2 to 10 species with $\sigma_c = 0.1$ and $l_{max} = n/2$. Red intervals show results for uncorrelated model as described in the legend to Fig. 3, blue intervals show results for model with eigenvalues $n > l_{max}$ coerced to zero, as described in Appendix D.I in the supplement, and purple regions show overlap between the two.

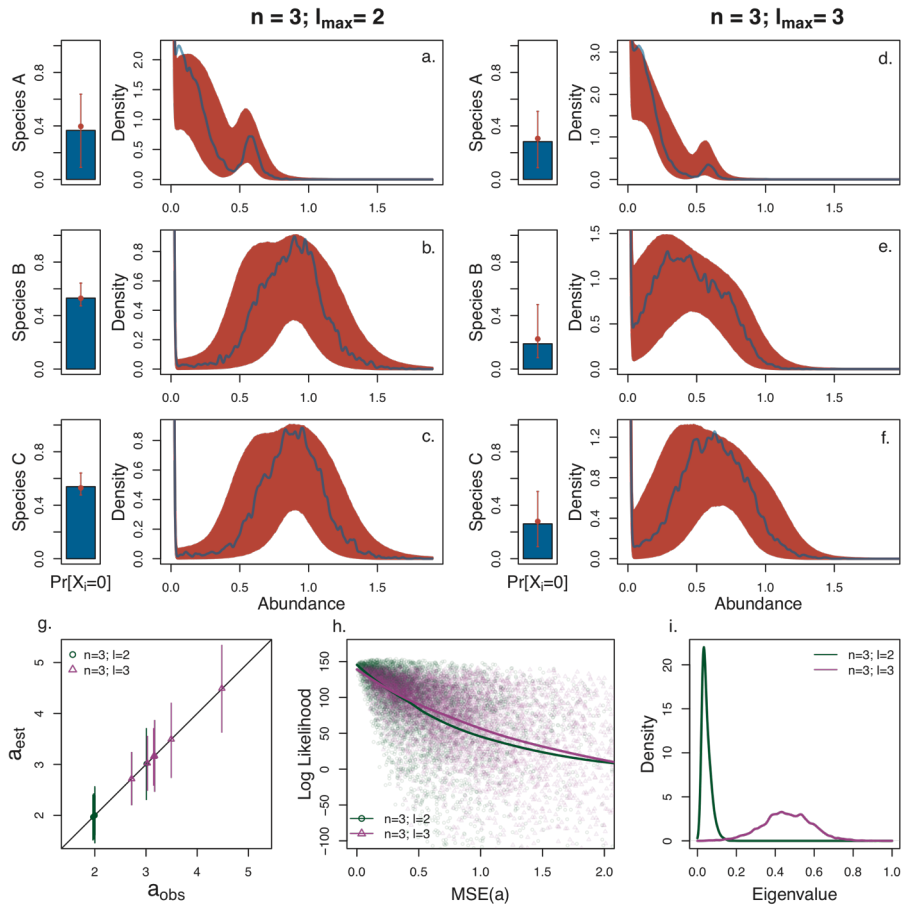

Figure 5: Fitting the covariance model to empirical data. Figures show variability in predictions from the covariance model based on observations from simulations from the mechanistic model of 100 "plots". **(a-f)** Frequency distributions of species abundance for a system with $n=3$ and $l_{max}=2$ **(a-c)**, and for a system with $n=3$ and $l_{max}=3$ **(d-f)**. Bar plots show frequency of zero abundances for each species, and density plots show distribution of positive abundances. Blue shows expected results for the mechanistic model, while red shows mean ± one standard deviation based on results from 5,000 iterations of the covariance model. **(g)** Comparison of observed and estimated elements of $\boldsymbol{a}$. Vertical lines show 95% confidence intervals for parameter estimates from the covariance model. **(h)** Model likelihood as a function of the mean square error (MSE) comparing estimated and observed parameters for the interaction matrix, $\boldsymbol{a}$. **(i)** Distribution of the third (smallest) eigenvalue of $\boldsymbol{a}$ for the two systems.

Chapter Three:

Diversity, high dimensional tradeoffs, and trait evolution buffer effects of reduced resource limitation on coexistence

Abstract:

Substantial evidence shows that increasing the availability of limiting resources can cause competitive exclusion among previously coexisting species. Theory suggests that these results occur because of tradeoffs, which require that increased competitive ability for a particular resource necessarily corresponds to decreased competitive ability for others. Here, we explore how these effects depend on initial community diversity and the number of resources by which species are limited, using a model of community assembly and evolution among species competing for multiple resources. Matching existing results, our model predicts that investing heavily in acquiring a particular resource leaves species vulnerable to competitive exclusion when that resource is no longer limiting. Subsequently, ecological and evolutionary selection drive communities towards the lower-dimensional "tradeoff frontier" defined by the remaining resources. Our model provides new evidence that diverse communities approach this frontier more rapidly, and initial differences among species have smaller effects in systems with a greater number of limiting resources. Our results therefore suggest that even in diverse systems with high-dimensional tradeoffs, it may be possible to identify low-dimensional indicators of coexistence mechanisms. These findings could be particularly useful for predicting, or even preventing, extinction caused by factors such as atmospheric nutrient deposition or landscape homogenization.

Introduction:

Tradeoffs are important for both empirical and theoretical understanding in ecology and evolution. Broadly, tradeoffs describe alternative strategies among which increased investment in one necessarily requires decreased investment in others (Stearns 1989; Zera & Harshman 2001; Roff & Fairbairn 2007). Most ecological theory predicts that stable coexistence requires that all species adhere to the same tradeoffs (Tilman

1990, 2011; Chesson 2000), as otherwise a single species could come to competitively displace all others. Tradeoffs are therefore important both in determining which species are able to coexist (Tilman 1990, 2011), and in structuring how they evolve in response to ecological and environmental changes (Stearns 1989).

Tradeoffs are often invoked to explain observed reductions in diversity following experimental addition of resources. This result is supported by substantial empirical evidence, primarily from grassland plant communities (Tilman 1982, 1987; Harpole & Tilman 2007; Hillebrand *et al.* 2007; Hautier, Niklaus & Hector 2009; Isbell *et al.* 2013; Harpole *et al.* 2016). In general, it is thought that adding sufficient quantities of a previously limiting resource causes it to no longer limit species growth, leading other resources to become limiting instead. Tradeoffs imply that species that have invested heavily in competitive ability for the added resource consequently invested less heavily in competitive ability for other resources, and are therefore likely to be competitively excluded when they become limited by them (Tilman 1987, 1990).

In reality, incorporating empirical results into theoretical understanding of tradeoffs is somewhat more complicated. Observed interspecific relationships among traits, as well as species responses to changes such as removal of a previously limiting factor, are jointly influenced by both "physiological" and "ecological" tradeoffs (Tilman 1990, 2004). Physiological tradeoffs constrain traits of individual organisms based on unavoidable structural and energetic limitations. For example, a plant cannot simultaneously allocate the same unit of resources to both roots and leaves (Gleeson & Tilman 1990). Conversely, ecological tradeoffs arise from interactions among organisms, or between organisms and their environment, and limit species to a subset of relationships among traits that allow for coexistence. For example, competition and limiting similarity might drive species with traits that are too similar to one another extinct, even if these traits are physically possible (Hutchinson 1959; MacArthur & Levins 1967).

A combination of many factors determines how these two types of tradeoffs influence coexistence, including the number of species in a community, the total number of limiting resources, species evolutionary histories, the distribution of resources across space and time, the types of interactions among species, and even stochastic events

(Tilman 1982, 1990, 2004, 2011; Lehman 2000; Chase & Leibold 2003; Farrior *et al.* 2013). In particular, results from existing theoretical studies suggest that reductions in the number of limiting resources should have the greatest impact on communities that are diverse (Tilman 2004), have few other limiting resources (Tilman 1982), and include species that have evolved under different environmental circumstances (Tilman 2011).

Here, we test how these three factors jointly alter the effects of resource addition on ecological communities. To do so, we use a theoretical model of resource competition in a heterogeneous environment. Specifically, we: *(1)* generate models for systems limited by two or three resources; *(2)* populate these models with communities of various numbers of species that share the same physiological tradeoffs; *(3)* allow these communities to evolve until they reach stable trait distributions; and *(4)* test how communities respond when a limiting resource is made super-abundant, and when invaded by species from a different community with an evolutionary history in the lower-dimensional environment. We also generalize our results to larger communities with up to five limiting resources, and compare them to results from resource addition experiments in grasslands.

Methods:

Model Structure

We utilize a resource competition model with consumers and "abiotic resources" *sensu* Armstrong and McGehee (1980), which is a simple linearization of the model of Tilman 1976, 1977, and has been previously suggested as a case of resource competition that is isomorphic to classical Lotka-Volterra competition (Volterra 1926; MacArthur 1970; May & MacArthur 1972; Tilman 1976, 1977, 1982). In this model, R_j is the concentration of resource j, $R_{max,j}$ is the concentration of the resource in the absence of consuming species (i.e. "resource availability"), N_i is the abundance of species i, $R_{i,j}^*$ is the lowest concentration of resource j at which species i has positive growth, $r_{i,j}$ describes the increase in growth rate of species i per unit of resource j, and $q_{i,j}$ is the tissue concentration of resource j in species i:

$$dN_i/dt\ (1/N_i) = \min_j[r_{i,j}(R_j - R_{i,j}^*)] \qquad \text{Eq. (1a)}$$

$$R_j = R_{max,j} - \Sigma_i(q_{i,j}N_i) \quad \text{Eq. (1b)}$$

For simplicity, we assume that species share the same growth rate response for all resources (specifically, for simplicity we assume $r_{i,j} = 1$ for all i and j), and that species forage optimally for resource (i.e. $q_{i,j} = w_i R_{i,j}^*$, where w_i is a scaling constant; specifically, for simplicity we assume w_i=1 for all i). See *Appendix A.I* in the supplement for specific details of model derivation.

In this model, species with lower $R_{i,j}^*$ are able to persist on lower concentrations of R_j, and are therefore "superior competitors" for that resource, as they can draw resource concentrations down to levels that exclude species with higher R_j^*. In the absence of competition, carrying capacity for species i, K_i, depends on local resource availability as:

$$K_i = \min[(R_{max,j}-R_{i,j}^*)/R_{i,j}^*] \quad \text{Eq. (2)}$$

Coexistence criteria

For a two-species system, interaction strengths can be derived from Eq. (1a-b) as:

$$\alpha_{A,B} = R_{B,j'}^*/R_{A,j'}^* \quad \text{Eq. (3)}$$

where A and B are two competing species, $\alpha_{A,B}$ is the per-capita effect of species B on the growth rate of species A, and $R_{A,j'}^*$ and $R_{B,j'}^*$ correspond to resource $R_{j'}$, which is the resource that is most limiting for species A. $R_{j'}$ can be identified by calculating growth rate for species A following Eq. (1a-b), and is the resource corresponding to the lowest growth rate.

Following standard criteria for Lotka-Volterra competition *sensu* MacArthur (1972), species stably coexist only if the following two constraints are met:

$$K_A/\alpha_{A,B} > K_B; \text{ i.e. } (R_{max,X}-R_{A,X}^*) > (R_{max,Y}-R_{B,Y}^*)R_{B,X}^*/R_{B,Y}^* \quad \text{Eq. (4a)}$$

$$K_B/\alpha_{B,A} > K_A; \text{ i.e. } (R_{max,Y}-R_{B,Y}^*) > (R_{max,X}-R_{A,X}^*)R_{A,Y}^*/R_{A,X}^* \quad \text{Eq. (4b)}$$

where species A is most limited by R_X, and species B is most limited by R_Y. These criteria can only be satisfied when there are interspecific tradeoffs and both species are limited by the resource for which they are an inferior competitor (i.e. $R_{A,X}^* > R_{B,X}^*$, and $R_{B,Y}^* > R_{A,Y}$) (Tilman 1982 pp. 190-204).

At the "local" scale (i.e. a single site), this model does not allow the number of coexisting species to exceed the number of limiting resources (MacArthur & Levins 1964; Levin 1970). However, given a spatial gradient along which resource availabilities vary, any number of species can potentially coexist with as few as two limiting resources (Tilman 1982). A necessary criterion for this high-dimensional coexistence is that species must all be subject to the same physiological tradeoffs (as defined above), such that: *(i)* increased competitive ability for one resource corresponds to decreased competitive ability for another, and *(ii)* relationships among R^* values must be shared across all species (Gleeson & Tilman 1990; Tilman 1990, 2011).

Given sufficient heterogeneity that all potential ratios of resource availability exist, then any species that falls along this physiological tradeoff should be able to either competitively exclude or coexist with any other combination of species in at least a subset of sites. Alternatively, if heterogeneity is limited, then some trait combinations will lead to competitive exclusion from all sites. Thus, the ecological tradeoffs in this system are jointly determined by the total number of limiting resources, and the level of resource heterogeneity (Tilman 1982).

Evolution and Community Assembly

To model ecological communities with evolutionary histories in different kinds of environments, we simulated systems with two limiting resources (R_X and R_Y), and with three limiting resources (R_X, R_Y, and R_Z). For simplicity, we refer to environments with three limiting resources as $R_{X,Y,Z}$, and those where only X and Y are limiting as $R_{X,Y}$. Similarly, we refer to groups of species that have evolved in $R_{X,Y}$ environments as N_2 communities, and those that evolved in $R_{X,Y,Z}$ environments as N_3 communities.

A simple way to meet the requirements outlined above for physiological tradeoffs is to assume linear, negative relationships among species competitive abilities for each resource, and to draw R^* values for each species from a single, shared surface. As such, to define physiological tradeoffs in our simulations, we constrained species traits to fall along a three dimensional plane, of the form $R_{i,X}^* + R_{i,Y}^* + R_{i,Z}^* - 2 = 0$ (note that decreasing any one R^* requires increasing other R^* values). We also include the

restriction $0 < R_{i,j}^* < 1$. Though the specific range of the restriction does not qualitatively influence results, its general purpose is to prevent species from sacrificing an infinite amount of competitive ability for one resource in order to gain infinite competitive ability for another (which is physiologically unrealistic).

For simulations of N_2 communities, we fixed $R_{i,Z}^*$ at a value of 1 (i.e. the "optimum" trait value for $R_{X,Y}$ environments), and sampled $R_{i,X}^*$ and $R_{i,Y}^*$ for species using random, uniform draws from the resulting line $R_{i,X}^* + R_{i,Y}^* - 1 = 0$. We then placed these species within $R_{X,Y}$ environments, each containing 1000 sites with resource availability varying across evenly spaced values on the line $R_{X,max} + R_{Y,max} - 2 = 0$ (we assumed that $R_{Z,max}$ was sufficiently large that it was never limiting). For simulations of N_3 communities, we chose species traits by sampling random, uniform draws from across the full physiological tradeoff plane. We then placed these species into $R_{X,Y,Z}$ environments, which were made up of sites spanning a fully factorial combination of all three resources using 100 evenly spaced levels for each resource (i.e. 5044 sites in total), with the restriction $R_{X,max} + R_{Y,max} + R_{Z,max} - 4 = 0$. These gradients thus yielded sites with all possible ratios of the two or three resources, respectively (within minor limits caused by discretization along the gradient). This means that all in N_2 species could persist somewhere along the $R_{X,Y}$ and $R_{X,Y,Z}$ gradients, and all N_3 species could persist somewhere along the $R_{X,Y,Z}$ gradient (including some sites in which no N_2 species could persist in monoculture). Note that the tradeoffs and resource gradients described here match those in Figs. 1-2.

For N_2 communities, we calculated predicted equilibrium species abundances in each site using Eqs. (2-3), and the stability criteria in Eqs. (4a-b). For N_3 communities, rather than calculating the stability of all potential equilibria for all sites (of which there could be more than 10^{15} per site for a community of 50 species), we computationally integrated Eqs. (1a-b) for the subset of species with positive carrying capacities, K_i. For both models, this process generated a single estimated equilibrium abundance for each species in each site (see *Appendix C* for details).

To represent evolution, we allowed species to take random "walks" in trait space, bound to the physiological tradeoff, following a random uniform distribution within a

fixed distance from their original R^* values (0.1 units for N_2 communities and 0.2 for N_3 communities; larger step size accounted for increased dimensionality). If this new combination of traits increased that species' "global" abundance (i.e. summed abundance across all sites), it took on these new traits, whereas if it did not, then a species retained its initial traits. We used this method iteratively across each species, alternating among species each time step. We use this method because increases in global abundance indicate a temporary increase in the population growth rate, which should be favored by natural selection. Though it assumes that variation arises in populations at a constant rate, similar abstractions are common in many community phylogenetic methods (Kraft *et al.* 2007; Cavender-Bares *et al.* 2009), and in existing ecological models of trait evolution in diverse communities (Rangel, Diniz-Filho & Colwell 2007; Gotelli *et al.* 2009).

For both the two- and three-dimensional models, we simulated system dynamics through time to determine the stable patterns of species abundance and trait distributions that arose from this model (the "evolutionarily stable configuration"). Preliminary testing suggested that 1,000 time steps was a sufficient "burn-in" period to generate the evolutionarily stable community configuration (also see example in Fig. 3d-e). To quantify variability in the evolutionary stable configuration, we ran simulations for 100,000 time steps, recording species abundance and traits every 100 time steps after the initial burn-in period.

Invasion and Competitive Exclusion

Next, we used the models described above to test whether species that adhered to the same physiological tradeoffs, but evolved under different conditions, could invade one another's established communities. To simulate this process, we generated evolutionarily stable configurations for N_3 communities, placed them in $R_{X,Y}$ environments, and invaded the resulting community with species from evolutionarily stable configurations for N_2 (n.b. all species from the N_3 community could persist in monoculture in the $R_{X,Y}$ environment – see *Appendix A.II* in the supplement).

Because more diverse communities have been shown to be more resistant to invasion (Tilman 2004; Fargione & Tilman 2005), we tested four scenarios with resident

N_3 communities of 5, 10, 25, and 50 species. Invading community richness was held constant at 10 for all scenarios. For each scenario, we simulated 1,000 independent communities and invasion events, and recorded invader success and abundance, as well as the abundance and survival of species from the original resident community. To assess the resident community's ability to adapt to new local conditions, we also tracked trait evolution in N_3 communities after their transition to $R_{X,Y}$ environments in the absence of invasion.

Lastly, we used a simplified set of methods to test results for larger communities with larger numbers of limiting resources, because simulating dynamics for these systems rapidly becomes computationally infeasible. For systems of between 2 and 5 limiting resources, and between 5 and 100 resident species, we sampled traits from the physiological tradeoff $R_{i,1}^* + R_{i,2}^* + \ldots + R_{i,m}^* - (m-1) = 0$, where m was the number of limiting resources (thus, N_2 and N_3 communities are subsets of this higher-dimensional surface). We then calculated mean fraction of species that were predicted to be universally inferior competitors if one of the limiting resources was removed (i.e. all R^* values were greater than that of another species in the community), based on 20,000 simulations. Lastly, we "invaded" these communities with species drawn from physiological tradeoffs with $m-1$ dimensions (i.e. including all resources but the one assumed to be no longer limiting), with invading community sizes ranging from between 5 and 50 species, and again estimated mean competitive displacement.

Grassland Nitrogen Addition Experiment

To test whether model results matched empirical evidence, we compared them to grassland plant species abundance and trait distributions from a long-term nitrogen addition experiment at the Cedar Creek Ecosystem Science Reserve (Cedar Creek) in Minnesota, USA. Cedar Creek is strongly nitrogen limited, and thus, nitrogen fertilization removes a previously limiting resource, similar to our simulated scenarios (Tilman 1987). These plots were established in abandoned agricultural fields in 1982, and have received fertilization treatments of between 0 and 27.2 g nitrogen (N) $m^{-2}yr^{-1}$. Except for a "no-treatment" control, plots also received annual addition of micronutrients to ensure that

nitrogen was the only limiting soil nutrient. Full methods are described in Tilman (1987). We used data from 54 plots in a single field ("C"), which was abandoned from agricultural use in 1934. To exclude establishment effects and later changes in experimental design, we used only data collected between 1990 and 2004.

For 24 of these species, we obtained measurements of R^* for soil nitrate (R_N^*) from monocultures grown as part of other experiments at Cedar Creek (see Table S1 in the supplement). In previous experiments, R_N^* has been found to be an effective indicator of competitive hierarchy for nitrogen (Wedin & Tilman 1993). These 24 species accounted for between 12% and 39% of total biomass in the experimental nitrogen addition plots (mean ± one standard deviation), and include 11 of the 20 most common species at Cedar Creek (including the four most common species). For each species and nitrogen treatment, we tested associations between abundance and R_N^*. Because we found no difference between "no treatment" and "control treatments" (i.e. with and without added micronutrients), we analyzed both of these treatments as having received 0 g N $m^{-2}yr^{-1}$.

Results:

Evolution and Community Assembly

After the burn-in period, we found that communities settled at stable abundance and trait distributions for both N_2 and N_3 communities. However, though species reached stable, positive equilibrium population sizes, they did not evolve to have identical traits or equal abundances. For example, in N_2 communities, species followed a classical abundance distribution, with a small number of common species, and a large number of rare species (Fig. 3a). Species with intermediate values of R_X^* and R_Y^* (hereafter "generalists", because they are intermediate competitors for multiple resources) tended to be the most common, while species with extreme trait values (e.g. $R_X^* \approx 1$, $R_Y^* \approx 0$, hereafter "specialists", because they are a good competitor for a single resource, but a poorer competitor for other resources) tended to be rarer (Fig. 3b). Species frequency distributions followed the opposite trend: while relatively few species evolved to be generalists, many species evolved to be specialists (Fig. 3c).

Invasion and Competitive Exclusion

In all cases where we reduced the number of limiting resources from three to two, more than half of the species from N_3 communities were competitively excluded (Fig. 4a-b,d). The surviving species were grouped near the $R_{X,Y}$ "tradeoff frontier" (i.e. $R_Z^*=1$), though species never fell precisely on the frontier itself (Fig. 4b). As resident community richness grew, species packed more tightly in trait space (Fig. 4c), and the distance between resident species traits and the fitness frontier shrank. Trait packing was denser for all but the least diverse resident community than it was for the invading community.

Following invasion, resident species abundance and survival were strongly associated with R_Z^*, such that species with higher R_Z^* (i.e. species that were closer to the two-dimensional tradeoff frontier) had higher abundance (Fig. 4e) and lower probability of extinction (Fig. 4f). In more diverse communities, more resident species initially had traits near the frontier, and consequently species that were farther from the tradeoff frontier tended to have somewhat lower relative abundance and higher extinction probability (Fig. 4e-f). When allowed to evolve after the transition from $R_{X,Y}$ to $R_{X,Y,Z}$ environments in the absence of invaders, N_3 communities rapidly evolved towards the $R_{X,Y}$ tradeoff frontier (Fig. 3d-e). Even in the absence of invasion, trait evolution in the N_3 community quickly led to competitive displacement of most resident species.

Though more resident species were displaced by other resident species than were displaced by invaders, invading species from R_2 communities established and competitively excluded some resident species in all simulations. Exclusions caused by resident species increased with resident species diversity, while exclusion caused by invaders declined (Fig. 4d). These results were largely consistent with our approximations for larger communities and communities with more limiting resources (Fig. 5). In these analyses, competitive exclusion caused by invading species also declined as a function of initial resident community diversity, though this effect grew weaker as the size of the invading community increased. Exclusions of all forms were rarer in communities with more limiting resources.

Grassland Nitrogen Addition Experiment

In control plots and at low nitrogen addition levels, there was a monotonic, negative relationship between species abundance and R_N^*, such that the most common species were those with lowest R_N^* (Fig. 6a-b). At higher nitrogen addition rates, abundance of species with low R_N^* decreased, and abundance of species with intermediate R_N^* increased, leading to a "hump-shaped" pattern. Interestingly, the pattern in high nitrogen plots qualitatively matched the expected evolutionary stable distribution of traits observed in systems with multiple limiting resources (Fig. 3b), whereas that in low-nitrogen plots more closely matched expectations from systems with a single limiting resource, with highest abundances among the best nitrogen competitors. Similarly, the distribution of species traits remained consistent across nitrogen addition treatments, with a majority of species having lower values of R_N^* (Fig. 6e-h).

Discussion:

Our results demonstrate several important aspects of how physiological and ecological tradeoffs structure community assembly and evolution, and how they influence community responses to resource addition. First, we find that evolution along physiological tradeoffs is strongly dependent on ecological interactions among species, and that these interactions lead to divergence both in terms of traits and species abundances. Second, we show that removal of a previously limiting factor can lead previously coexisting species to be competitively excluded both by other resident species, and by newly arrived invading species from habitats in which the previously limiting factor had always been in high abundance. Third, we find that both initial community size and number of limiting factors somewhat buffer the effects of removing a limiting factor, though not always in ways that reduce species loss. Lastly, our results show that ecological and evolutionary forces quickly drive species towards new tradeoff frontiers after a limiting factor is removed, suggesting that the observed "dimensionality" of communities may be more strongly determined by local factors than by high-dimensional physiological tradeoffs (see *Diversity and dimensionality of observed tradeoffs* below).

Evolution and Community Assembly

It may seem counterintuitive that species, which all have the potential to evolve towards the same set of traits, do not form communities in which species are equally abundant, which is a feature that earlier work also found (Tilman 2004). This result occurs in our model because of asymmetrical competitive interactions between "generalists" (i.e. species that are intermediate competitors for all resources) and "specialists" (i.e. strong competitors for one resource but poor competitors for others). For example, consider a specialist, species *A* with $R_{A,X}^* = 0.85$ and $R_{A,Y}^* = 0.15$, competing against a generalist species *B* with $R_{B,X}^* = 0.45$ and $R_{B,Y}^* = 0.55$ (illustrated in Fig. 1b,d). Local competitive interactions will yield $\alpha_{A,B} = R_{B,X}^*/R_{A,X}^* = 0.529$, and $\alpha_{B,A} = R_{A,Y}^*/R_{B,Y}^* = 0.273$ (n.b. figure shows α for global competitive interactions), meaning that the effect of the generalist on the specialist is greater than is true for the reverse interaction.

More broadly, it is possible to show that competitive effects of generalists on specialists will always be stronger than the reverse for any two species in our model (see *Appendix B.I* in the supplement). This competitive asymmetry has major implications for how species evolve, because changes in traits that make generalists more like specialists reduce abundance of the specialist more than they reduce abundance of the generalist (n.b. because competitive effects are additive across species in our model, the same general result hold for communities of any size). In any community that includes multiple competing species, selection will therefore drive some generalist towards the edges of the physiological tradeoff, because this change reduces competitive impacts from other generalists more than it increases competitive impacts from specialists. This phenomenon is effectively a directional version of ecological "squeeze" (MacArthur & Levins 1964; MacArthur 1972), and matches a well-documented empirical pattern known as the taxon cycle (Wilson 1961; Ricklefs & Bermingham 2002; Tilman 2004). Similar results have been demonstrated in other theoretical models of evolution along tradeoffs as well (Lehman 2000).

Importantly, the direction of competitive asymmetry varies depending on the precise structure of the model. For example, in models of competition for "biotic"

resources *sensu* Tilman (1982) and Armstrong and McGehee (1980), the same results as above hold for essential resources, whereas specialists are favored in competition for perfectly substitutable resources, and either specialists or generalists can be favored in competition for switching resources, depending on local conditions (see *Appendix B.II* in the supplement) (Tilman 1982, 2004). These variable results suggest that depending on the types of resources for which species compete, evolutionary and ecological selection may drive species towards different regions in physiological tradeoff space. These differences could be useful indicators for detecting competitive mechanisms based on species trait distributions (Tilman 2004).

Competitive Exclusion and Invasion

Our theoretical results showing steep declines in diversity following removal of a limiting factor also are consistent with observed declines in diversity following nitrogen fertilization, and match previously published theory on resource competition (Tilman 1982, 1987; Hautier *et al.* 2009). Our findings are also consistent with experimental results showing that each additional limiting resource added to a system leads to additional loss of species, and that more diverse communities lose a greater fraction of resident species following resource addition (Interlandi & Kilham 2001; Harpole & Tilman 2007; Harpole *et al.* 2016). Interestingly, empirical results suggest that species are lost either as a linear function of number of resources added, or potentially even at an increasing rate, matching our results for communities of between 5 and 25 species, but not for larger communities (Fig. 5a). Note that in our analyses, when community sizes are small (i.e. < 20 species), removing one of a small number of limiting resources (e.g. removing one of three resources) results in a larger change in the fraction of species lost than does removing one of a larger number of limiting resources (e.g. removing one of five), whereas when community sizes are large (e.g. 100 species), the reverse is true. This result occurs because in diverse communities in our model, a small number of species are sufficiently close to the tradeoff frontier to exclude most competitors after removal of a limiting resource, even if many other limiting resources remain. The absence of a similar trend in empirical results may suggest that communities in these experiments are

insufficiently diverse to fully saturate the range of traits related to local coexistence mechanisms.

Results from our invasion simulations accord with existing empirical evidence in several ways. First, higher diversity resident communities in our model were more effective at limiting invader abundance, matching commonly observed experimental results (Knops *et al.* 1999; Naeem *et al.* 2000; Fargione, Brown & Tilman 2003; Fargione & Tilman 2005). Second, across all scenarios, invaders persisted in established communities, and competitively excluded at least some resident species (Fig. 4c). This result accords with the predicted outcomes of "evolutionarily stable strategies", which suggest that resident species should be able to exclude invaders from a particular subset of conditions, but not from all potential conditions (Pimm & Rosenzweig 1981; Tilman 2011).

Our analytical model and simulation results suggest that two mechanisms are required for invading species to competitively exclude resident species which share the same physiological tradeoffs: *(i)* local environmental conditions must have diverged from the conditions under which resident species evolved, and *(ii)* invaders must be better adapted for the new environmental conditions than are resident species. These mechanisms are consistent with effects of experimental nutrient addition, which show that resident species that evolved in nutrient poor environments tend to be outcompeted by ruderal invaders when fertilization rates are high (Tilman 1987). These mechanisms also match general patterns showing higher rates of invasion following disturbance, particularly by species that have an evolutionary history in disturbed environments (Pysek *et al.* 2012).

Importantly, the spatial and temporal scales across which these two mechanisms can lead to extinction are likely limited, as selection should drive resident species towards lower-dimensional tradeoff frontiers in novel environments (e.g. Fig. 3d-e). This scale-dependency may explain why there are relatively few historical examples of competitive exclusion following invasion events across biogeographic regions (Tilman 2011). Similarly, in other systems with lower heterogeneity, such as those with limited niches

(Fargione *et al.* 2003) or that are subject to demographic stochasticity (Tilman 2004), invaders may be entirely excluded by sufficiently diverse resident communities.

Nevertheless, there are some cases where these mechanisms could be important contributors to extinction at smaller spatial and temporal scales. First, human-mediated invasions and changes in environments might be too fast for evolutionary responses to occur (Schlaepfer, Runge & Sherman 2002). Similarly, if local communities are particularly small, or species are highly specialized, then even small and gradual changes may be sufficient to exclude them from all sites that they can occupy (Davies, Margules & Lawrence 2004). Third, evolutionary constraints might prevent species from allocating resources in ways that optimally match changes in local conditions (Futuyma 2010). Lastly, heterogeneous conditions may lead species to invest in traits that are favored in some regions, but are less beneficial at local scales (Bridle & Vines 2007; Futuyma 2010). Broadly, these circumstances could all lead species to under-invest in competitive ability for particular resources, potentially allowing better-adapted competitors to exclude them from environments where those resources are scarce.

Diversity and dimensionality of observed tradeoffs

A potential problem implicated by our findings is that species have many ecologically relevant traits (Clark *et al.* 2010). If physiological and ecological tradeoffs are really conserved across all of these traits, then properly identifying these tradeoffs could require enormous amounts of high-dimensional data, which comes with many challenges. Most notably, these include the danger of erroneously attributing causal significance to spurious associations between community composition and particular coexistence mechanisms (Adler et al. 2013) or particular combinations of traits (Roff & Fairbairn 2007).

Luckily, interplays between physiological and ecological tradeoffs might structure evolution and community assembly in ways that facilitate identifying locally important coexistence mechanisms. At least at small scales, it appears that coexistence in many systems depends on only a modest number of limiting factors (Hutchinson 1959; Chapin 1980; Tilman 1982). Results from our models therefore suggest that selective assembly

and subsequent trait evolution should lead to communities of species that invest strongly in traits relating to locally important coexistence mechanisms, and invest very little in other traits. Thus, in communities where coexistence is determined by a small number of factors, relationships among species traits should correspond to equally low-dimensional patterns.

This property could greatly simplify identifying influential traits and corresponding coexistence mechanisms in real-world communities. Depending on which traits follow constrained relationships under which types of conditions, it may be possible to use empirical observations of trait-based tradeoff surfaces to infer the identity of some of these mechanisms (McGill *et al.* 2006; Roff & Fairbairn 2007; Kraft, Godoy & Levine 2015b). Small-scale matching between traits and environments may also explain why there are relatively few examples of general associations between species traits and large-scale environmental gradients (Shipley *et al.* 2016). Testing relationships between traits and ecological outcome without specific consideration of the mechanism by which that trait acts may therefore yield limited insight (Roff & Fairbairn 2007; Adler *et al.* 2013).

Grassland Nitrogen Addition Experiment

The mismatch between observed species abundance distributions at low fertilization rates and those predicted by our model potentially suggests that without nitrogen fertilization, nitrogen is sufficiently scarce that it constitutes the only limiting resource, with multi-species coexistence potentially explained by other mechanisms, such as a competition-colonization tradeoff and metapopulation dynamics (Lehman 2000). At higher fertilization rates, observations showing highest abundance among species with intermediate R_N^* matched those for simulated communities with two limiting resources (Fig. 3b). Similarly, species trait distributions roughly matched simulated results across fertilization treatments (Fig. 3a). Importantly, both results hold even when legumes are excluded, as these specie tend to have higher R_N^* values and are less responsive to competition for nitrogen (Fargione *et al.* 2003). These patterns are consistent with expectations from tradeoffs, as fertilization should only reduce abundance among species with low R_N^* if it is also correlated with poor competitive ability for other resources that

were not previously limiting (Tilman 1982), such as light (Hautier *et al.* 2009) or water (Farrior *et al.* 2013).

Caveats

Though there is reason to believe that our results may be generalizable to other models and systems, there are a number of limitations to the approaches we use. Most obviously, the competition model we utilize includes many simplifying assumptions, such a single, linearized growth rate response to all resources that is shared across species ($r_{i,j}$). Similarly, the modeling framework we apply does not include interactions among multiple trophic levels, nor does it incorporate interactions that are not mediated through resources (Vellend 2016). While similar models have been developed that include these kinds of factors, results often do not follow trivially from their simpler analogs (Chase & Leibold 2003; Tilman 2004; Hodapp *et al.* 2016).

Even in the context of resource competition within a single trophic level, several fruitful directions exist for further research. First, we utilize resource supply gradients that include all possible ratios of resources, and thus allow large numbers of species to coexist at equilibrium. In reality, it is likely that only subsets of these conditions exist, which could contribute to environmental filtering and greater competitive exclusion, thus leading to substantially different abundance and trait distributions (Tilman 2004; Kraft *et al.* 2015a). Alternatively, ratios of nutrients (such as N:P:K:Ca) may vary with soil depth, and species may have rooting depth patterns that maximize their abilities to acquire their optimal resource ratios. Second, though we test how differences in community diversity influence invasion, we do not include any specific mechanism for determining the size of resident and invading communities. Models that include more realistic mechanisms may lead to communities that are relatively saturated with respect to locally important traits, which could lead resident communities to be more resistant to the effects of invading species (Tilman 2004).

Conclusion

We find that the effects of removing previously limiting resources on ecological communities are strongly shaped by the number of species in the community, the initial number of limiting resources, and species evolutionary histories. While most of these results follow from existing theory about tradeoffs and coexistence (Tilman 1990, 2004, 2011), this is, to our knowledge, the first study in which all of these factors have been tested jointly. Furthermore, while most of our results are relatively straightforward in retrospect, they often diverge from simple expectations of how existing theory should expand to more complex circumstances. Taken together, our results suggest that future work which specifically links empirical knowledge of physiological tradeoffs with theoretical understanding of how coexistence mechanisms are structured by ecological tradeoffs will be necessary for developing better understanding of species assembly and evolution in diverse ecological communities.

Acknowledgments:

We are grateful to E. Borer, G. Fury, C. Lehman, C. Neuhauser, D. Renard, K. Thompson, D. Tilman, and D. Williams for helpful comments on earlier drafts of this manuscript. We also thank C. Lehman and S. Loberg for helpful discussions on the modeling framework that we employ. This work was supported by grants from the US National Science Foundation Long-Term Ecological Research Program (LTER) including DEB-8114302, DEB-8811884, DEB-9411972, DEB-0080382, DEB-0620652, and DEB-1234162. Further support was provided by the Cedar Creek Ecosystem Science Reserve and the University of Minnesota. A.T.C. was supported by an NSF Graduate Research Fellowship, base award number 00006595. Computing time was provided by the University of Minnesota Supercomputing Institute.

Figures:

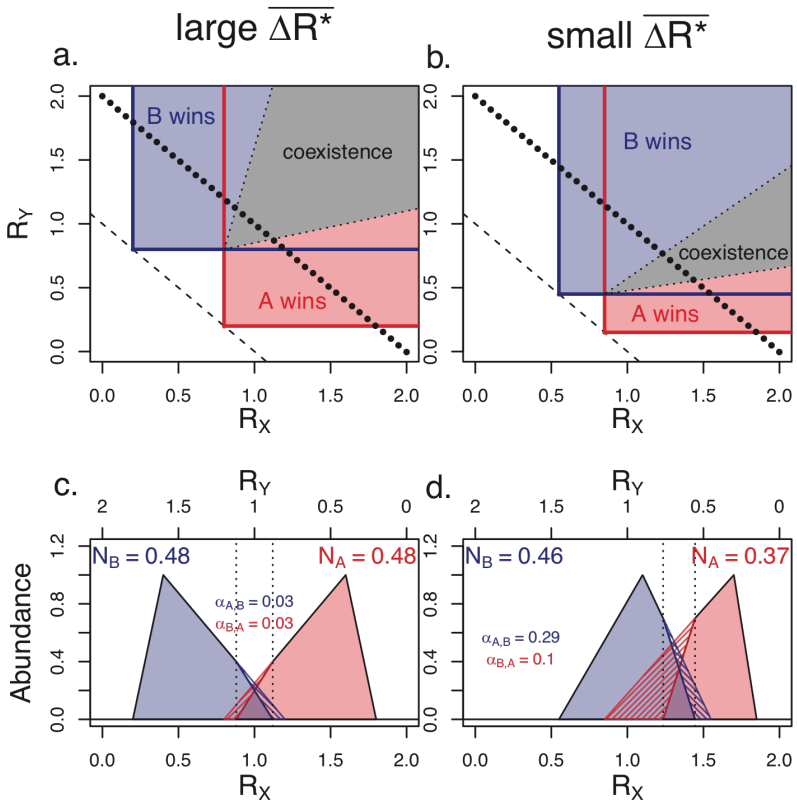

Figure 1: Competition between two species (*A* and *B*), both subject to the same physiological tradeoff in competitive ability between two resources (R_X and R_Y). **(a-b)** Conditions required for coexistence and competitive displacement. Solid red and blue lines show the zero net growth isoclines for species *A* and *B*, respectively. Dashed line shows the physiological tradeoff, such that R_X^* and R_Y^* must fall somewhere along the line. Black circles show a gradient in resource availability. Shaded regions show regions where species *A* competitively excludes species *B* (red), *B* excludes *A* (blue), or where they coexist (grey). **(c-d)** Species abundances corresponding to species pairs in the upper panels. Shaded regions show species abundance after accounting for competitive interactions (N_A and N_B), while hatched regions show abundance that is lost to competitive interactions. Note that K=0.5 for both species in all panels. $\alpha_{A,B}$ and $\alpha_{B,A}$ show "global" competitive effect of species *B* on *A*, and of *A* on *B*, respectively.

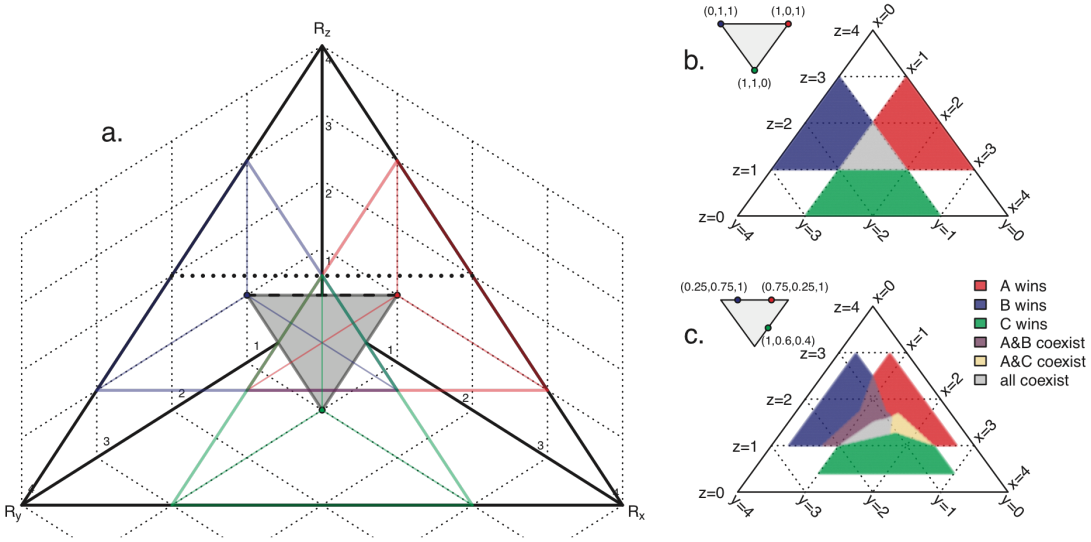

Figure 2: Competition among three species (*A*, *B*, and *C*, shown as red, blue, and green, respectively), subject to a shared physiological tradeoff in competitive ability for three resources (R_X, R_Y, and R_Z). **(a)** Three-dimensional rendering of physiological tradeoffs and resource gradients. Axes meet at R_X=R_Y=R_Z=0, similar to the corner of a room where the walls meet the floor. Grey triangle shows the physiological tradeoff plane; large black triangle shows resource availability gradient. Dashed line along the top of the tradeoff plane, and black circles on the resource gradient, correspond to the tradeoff and gradient in Fig. 1. Points show traits for each species; lines connect points to colored triangles on resource gradient, indicating regions where each species is able to persist. **(b-c)** Simplified renderings of the three-dimensional figure. Triangles in the upper left corner of each panel correspond to the tradeoff plane; large, black triangles in the center of each panel correspond to the resource gradient. Colored points indicate species traits; shaded areas indicate regions where species persist, coexist, or competitively exclude one another. Note that **(b)** corresponds exactly to **(a)**, while **(c)** shows species with different traits.

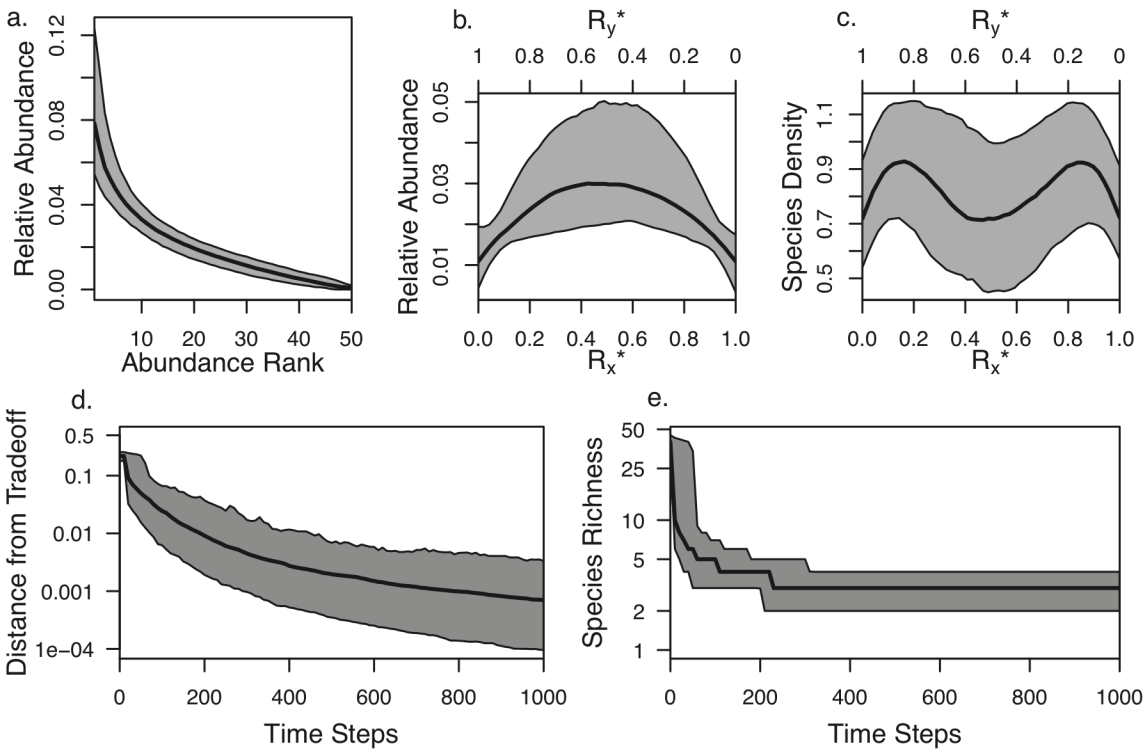

Figure 3: Evolution along a tradeoff in competitive ability for two limiting resources. Lines and shaded regions show median and 95% confidence interval, based on communities of 50 species. **(a-c)** Stationary trait distributions arising after 100,000 time steps. **(a)** Species abundance distribution. Though species are all subject to the same tradeoffs, they do not evolve to have identical abundances or traits. **(b)** Species abundances as a function of their R_X^* and R_Y^* (note that $R_Y^* = 1 - R_X^*$). "Generalists" (species with intermediate competitive abilities for both resources) are the most abundant, while "specialists" (good competitors for one resource, but bad competitors for the other) are rare. **(c)** Frequency distribution of species as a function of R_X^* and R_Y^*. Though they are rarer, there are more specialist than generalist species. **(d-e)** Evolution of species that originated in a region with three limiting resources, but must persist in a region where only two of those resources are limiting. Over time, species traits approach the two-resource tradeoff frontier **(d)**, though in the process, many are competitively excluded **(e)**.

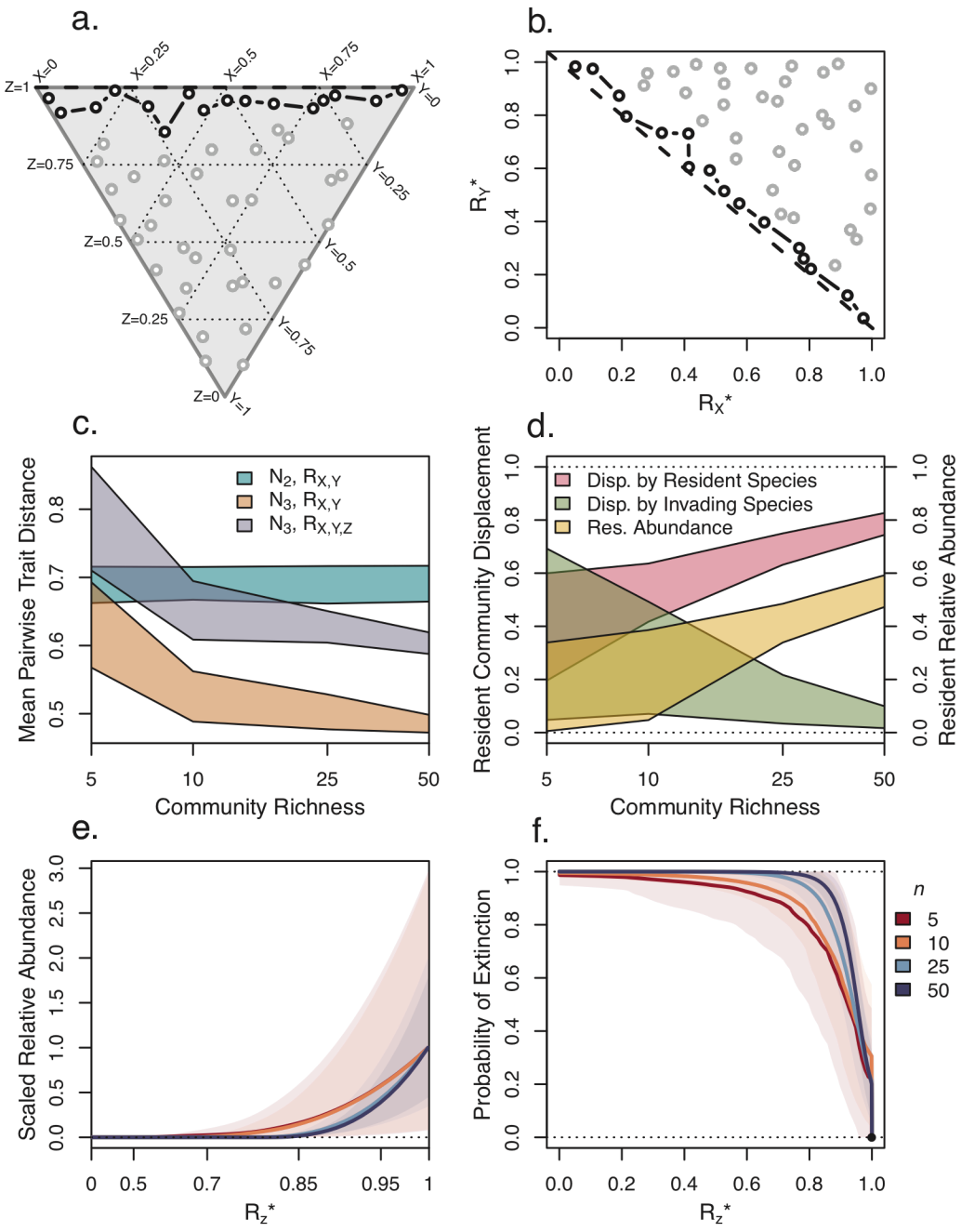

Figure 4: Evolution and invasion as a function of the number of limiting resources. **(a-b)** Example of a community of species that have evolved in a region with three limiting resources, after one of those resources (R_Z) ceases to be limiting, with traits shown in **(a)** three and **(b)** two dimensions. Dashed line shows the tradeoff frontier. Black points show species that survive under the new conditions; grey points show species that are out-competed. **(c-f)** Scenarios of species with different evolutionary histories invading into

established communities. “Community richness” (n) shows initial number of species in the resident community; invading community size is held constant at ten species. Resident species evolved in an environment with three limiting resources, but now inhabit an environment where R_Z is no longer limiting. Invading species evolved in an environment with only R_X and R_Y limiting. Shaded regions shows mean +/- one standard deviation, based on 1,000 simulations. **(c)** Mean Euclidian trait distance among resident species (orange and purple), or among invading species (blue). $R_{X,Y}$ refers to distance calculations based only on R_X and R_Y, while $R_{X,Y,Z}$ also includes R_Z. **(d)** Fraction of total resident community competitively displaced by other resident species (red), or by invading species (green). Yellow shows abundance of surviving resident species relative to invaders after invasion. **(e-f)** Abundance and extinction probability as a function of R_Z^* for resident species after the invasion event. For comparability across community sizes, mean abundances are scaled to have a maximum of one.

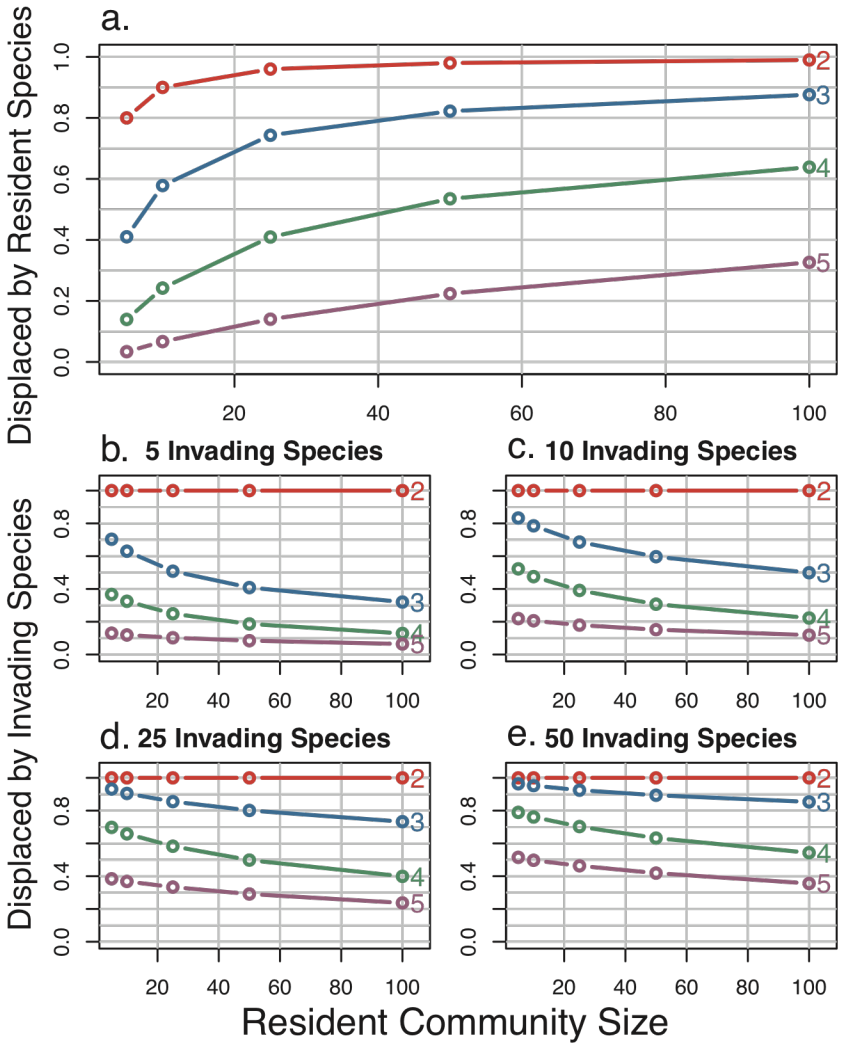

Figure 5: Estimated fraction of resident community competitively excluded following removal of one previously limiting resource. Numbers indicate how many resources were initially limiting; resident community size indicates initial number of resident species in the community. **(a)** Fraction of resident species competitively excluded by other resident species. Note that red line (originally 2 limiting resources) indicates that all but one species are competitively excluded. **(b-e)** Fraction of resident species competitively excluded by invading species, where traits for the invading species are drawn from the lower-dimensional tradeoff surface. Numbers above panels indicate the size of the invading community.

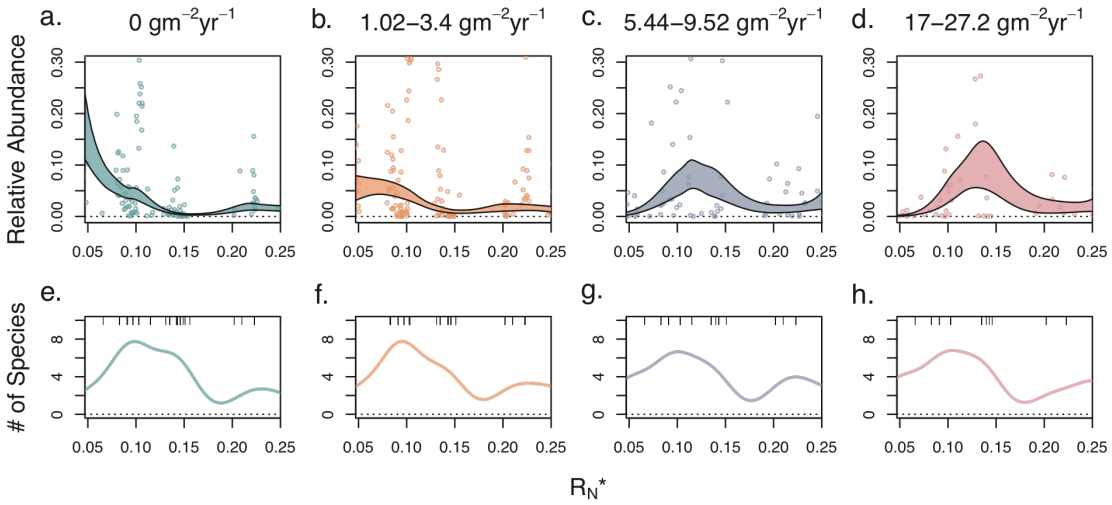

Figure 6: Species responses along an experimental nitrogen fertilization gradient at Cedar Creek, with addition rates of 0 g N $m^{-1}yr^{-1}$ **(a,e)**, 1.02-3.4 g N $m^{-1}yr^{-1}$ **(b,f)**, 5.44-9.52 g N $m^{-1}yr^{-1}$ **(c,g)**, 17-27.2 g N $m^{-1}yr^{-1}$ **(d,h)**. **(a-d)** Species relative abundance as a function of competitive ability for soil nitrate (R_N^*), comparable to Fig. 3b. At low fertilization rates, good nitrogen competitors (low R_N^*) are most abundant **(a-b)**. At higher fertilization rates, intermediate nitrogen competitors are most abundant **(c-d)**. Intervals show means +/- one standard error, based on loess smoother. **(e-h)** Distribution of species as a function of R_N^*, comparable to Fig. 3c. Ticks at top of panels show exact values for each species; colored lines show continuous estimates of frequency based on a kernel density smoother.

Chapter Four:

Tradeoff-Based Mechanisms Predict Coexistence, Productivity, and Species Abundances in Grassland Plant Communities

Abstract:

Ecological theory predicts that interspecific tradeoffs promote stable coexistence among competing species. Predictive ability of competition models may therefore be enhanced by explicit incorporation of empirically observed tradeoffs. We developed and tested such a model for perennial grassland plant species using traits related to nitrogen competition. As predicted by theory, observed trait measurements from monocultures of 35 species defined a distinct tradeoff surface. When parameterized with these traits, model predictions of species abundances and productivity were consistent with results observed in 122 multi-species mixtures of subsets of these species. Moreover, predictions improved markedly when observed species traits were "snapped" to the nearest location on the empirical tradeoff surface. Our results suggest that these traits and tradeoffs are important determinants of coexistence in our system. Furthermore, they show that models based on empirically observed interspecific tradeoffs could constitute useful and broadly applicable tests for identifying mechanisms that structure communities.

Introduction:

A primary goal of ecology is to understand the mechanisms that determine species abundances and diversity (Lawton 1999). Almost all known mechanisms that potentially explain observed patterns of diversity assume that species differ in how they interact with one another and their environments (Chesson 2000; Adler *et al.* 2013; May *et al.* 2016). Most theory therefore predicts that ecological communities are largely structured by the traits of species that determine these differences (McGill *et al.* 2006; Violle *et al.* 2007). Because most species traits are either unknown or imprecisely known, achieving this goal may require identifying subsets of traits and mechanisms that are best able to explain observed outcomes (Tilman 1990). However, it is not clear whether such subsets exist, nor how to identify them in any particular system.

Ecological tradeoffs offer a potential solution to this dilemma. Tradeoffs describe alternative strategies, among which increased investment in any one strategy requires decreased investment in others (Stearns 1989; Roff 1992; Zera & Harshman 2001; Roff & Fairbairn 2007). For example, by investing resources in roots, a plant necessarily forgoes – at least in the short term – potential allocations to leaves, stems, or seeds. For competing species to coexist, most theoretical frameworks require that all species be subject to the same tradeoffs (Tilman 1990; Chesson 2000), because any species with universally superior strategies would ultimately displace its competitors, and universally inferior competitors would not persist (Tilman 2011).

Different types of strategies are favoured depending on the ecological and evolutionary forces that have shaped communities, which can lead to distinct relationships among species traits (Roff & Fairbairn 2007). Relationships observed among particular subsets of traits could therefore indicate that these traits are related to tradeoffs that locally determine coexistence, and may help identify underlying coexistence mechanisms (Wright *et al.* 2004; Reich 2014; Kunstler *et al.* 2016). Furthermore, because the same tradeoffs must be shared across all coexisting species, this relationship is likely to be broadly generalizable, which could allow for accurate community-level predictions from models that are parameterized using only a subset of coexisting species (Purves & Pacala 2008; Litchman *et al.* 2012).

Ideally, relationships among traits that define tradeoffs should be derivable from basic biological assumptions (e.g. Brown *et al.* 2004). However, it is not clear that such a derivation is possible for many hypothesized coexistence mechanisms, which often include "high level" or abstracted traits. As an example, consider plant dispersal, which can depend on plant height (Thomson *et al.* 2011), seed physiology (Muller-Landau *et al.* 2008), seed size, biomass allocated to seed, abundances of seed predators (Janzen 1970; Connell 1971), meteorological conditions (Wright *et al.* 2008), or animal vectors (Harrison *et al.* 2013). Although many such selective forces likely guided the evolution of current day plant dispersal strategies, deriving a theory that simultaneously explains mechanistic tradeoffs among these factors seems unfeasible.

A useful alternative strategy has been to relate empirical observations of the relationships among species traits to theoretical tradeoffs among parameters in a model. For example, plant leaf biomass and area can relate to mechanistic parameters in models of light competition (Rüger *et al.* 2012; Farrior *et al.* 2013), and relationships among species foraging rates can be used to derive interaction coefficients in competition models (MacArthur 1970; Chesson 1990). Accurate predictions of species abundances from such models provide compelling evidence that their underlying coexistence mechanisms and tradeoffs could be important determinants of local community composition (Kneitel & Chase 2004; Litchman *et al.* 2012; Kraft & Ackerly 2014). However, minor mismatches between measured traits and the requirements for coexistence in the model can lead to erroneous predictions of coexistence or competitive exclusion (Tilman 2011). Even very small amounts of sampling noise or omission of some minor mechanisms could therefore obscure results, even if the model correctly characterizes a system's major mechanisms. Consequently, mechanism-based models often need to be "tuned" so that predictions better match observations (Pavlick *et al.* 2012; Kunstler *et al.* 2016; Evans *et al.* 2016). But once so tuned, it is unclear whether accurate predictions result from accurate mechanisms in the model, or from fortuitous fitting of an inaccurate model.

A solution proposed by Purves & Pacala (2008) is to assume that species traits fall exactly on the "tradeoff surface" characterized by empirically observed relationships among species traits, and thus to mathematically "snap" traits to the nearest corresponding point on the surface. Unlike methods that tune model parameters until a desired fit is achieved, snapping traits to empirical tradeoff surfaces merely smoothes across observed relationships among traits, based on the assumption that coexisting species have tradeoffs. If tradeoffs among these traits are primary determinants of coexistence, then parameterizing mechanism-based models using these snapped traits should reduce sampling noise and improve predictions. This is similar to approaches that "borrow power" from relationships observed across many species to better estimate trends for rare species (Evans *et al.* 2016).

We hypothesize that *(i)* if tradeoffs and mechanisms assumed in a model correctly describe major determinants of local coexistence, and *(ii)* if empirical tradeoffs among

traits corresponding to these mechanisms can be identified, then *(iii)* parameterizing models with traits snapped to this tradeoff surface should more accurately characterize underlying coexistence mechanisms and resultant species abundances than would models parameterized with "raw" unadjusted traits. We test this hypothesis using observations from long-term experiments in grassland plant communities. We first identify ecologically relevant traits based on prior observational and experimental studies in this system, and confirm that these traits fall along an empirical tradeoff surface. We then parameterize a mechanism-based model with either "raw" observed traits, or with traits "snapped" to this surface, and test for changes in the ability of these two parameterizations of this model to predict experimentally observed species abundances and coexistence for 122 different cases of multi-species competition. Our results suggest that formal inclusion of empirically observed tradeoffs in model parameterization improves model predictions. This method may therefore be useful for testing hypotheses about coexistence mechanisms, and for making more broadly generalizable species-level predictions in diverse ecological systems.

Materials and methods:

Site and Data Collection

To parameterize and test models and their corresponding tradeoffs, we used data from long-term grassland plant experiments at the Cedar Creek Ecosystem Science Reserve (Cedar Creek) in Minnesota, USA (45.40°N, 93.20°W). Soils at Cedar Creek are composed primarily of sandy glacial outwash and are strongly nitrogen limited (Tilman 1987). Mean annual precipitation is around 77 cm, with average daily summer high and winter low temperatures of 27°C and -14°C, respectively. We collected trait data from monocultures of 35 grassland plant species grown in seven experiments at Cedar Creek, for a total of 247 monoculture plots. These included 247 plots containing 35 grassland plant species. Three experiments (called E026, E055, and E070; cedarcreek.umn.edu/research/data) tested effects of soil fertility on competitive hierarchies; three (E120, E123, and E249) tested effects of planted diversity on ecosystem functioning; the last (E111) was a monoculture garden. All monoculture plots

were hand-weeded annually. We determined parameters for predictions with data only from "control" (i.e. un-manipulated) monoculture plots, collected at least three years after establishment. To test predictions, we used observations from 122 experimentally assembled multi-species plots. These 9m x 9m plots are part of the E120 biodiversity experiment at Cedar Creek. Plots were seeded in 1994 with randomly chosen mixtures of 2, 4, 8, or 16 species, and weeded annually to remove non-planted species (Tilman *et al.* 1997). Further details are available in Appendix *I* in the supplement.

Identifying influential traits

We focus on three characteristics measured in monocultures for a variety of species: competitive ability for soil nitrate, as measured by R^*, which is the measured per unit mass concentration of nitrate in monoculture soils; aboveground tissue nitrogen concentration q, measured as the percent of N in dry leaf mass; and aboveground biomass, B^*_{mono}, measured in g m^{-2} of dry biomass. Asterisks indicate that variables are assumed to be at equilibrium (i.e., no significant trend with plot age). B^*_{mono} describes total living biomass per square meter after drying to constant mass, measured with 10cm-wide clip strips; q was measured from homogenized subsets of biomass samples; R^* was estimated from soil cores taken at 0-20cm via 0.01M KCl extraction. Measurements were taken during "peak" biomass (late July or early August). For simplicity, we refer to these three characteristics as "traits", though they are measured at the population level and do not correspond to more commonly measured "functional traits" (Violle *et al.* 2007). Mean trait values and standard deviations were calculated for each species using mixed-effects regression models to account for differences in sample size among experiments (see Appendix *II* in the supplement for regression structures).

We use these traits because they relate to species abilities to acquire and retain nitrogen, which is a primary limiting resource at Cedar Creek (Tilman 1987). Species that drive soil nitrate to lower concentrations have lower R^* and are predicted to be superior nitrogen competitors (Wedin & Tilman 1993; Dybzinski & Tilman 2007) (see Appendix *III* in the supplement for examples). Species with lower q have higher nitrogen use efficiency because they produce more biomass per unit of aboveground tissue nitrogen

(Dybzinski & Tilman 2007). Finally, at fixed q, higher B^*_{mono} indicates that species access and retain a greater proportion of the total available nitrogen in aboveground tissues (Tilman 1994; Dybzinski & Tilman 2007). Note that we do not account for root biomass or its nitrogen content, which can be substantially greater than aboveground pools. We make this choice following the assumption that aboveground tissues are an indicator of annual fluxes, while roots represent longer-term averages of inter-annual growth, which can be more difficult to model.

To test the potential explanatory power of these traits, we fit a series of linear regressions to determine how mean species abundance in the 122 multi-species plots was statistically related to the three traits, and their two- and three-way interactions. We also estimated total community biomass as the sum of species abundances. Regressions included up to two additional covariates, which were interacted with all other terms in the model: all regressions included percent soil carbon, which is an indicator of nitrogen mineralization rates (Fornara & Tilman 2008); a subset of regressions included planted species richness, which strongly influences community productivity at Cedar Creek (Tilman *et al.* 1997). For all regressions, we log-transformed species abundance, percent soil carbon, planted richness, B^*_{mono}, and R^*, and logit-transformed q (see Appendix *IV* in the supplement for details on regression structure).

Empirically fitting the tradeoff surface

To test for tradeoffs among traits, we fit a surface to observed R^*, q, and B^*_{mono} using ranged major axis regression, which simultaneously minimizes errors in all variables. Again, we log-transformed R^* and B^*_{mono}, and logit-transformed q. This regression resulted in a two-dimensional plane in transformed parameter space. We then generated tradeoff-based trait predictions by projecting observed traits of each species to the nearest location on the fitted tradeoff surface, which we denote as $\hat{R}^*$, $\hat{q}$, and $\hat{B}^*_{mono}$. We refer to this process as "snapping" traits to the tradeoff surface, in contrast to "raw" trait values observed in monocultures. Detailed methods for the regression and snapping procedures are in Appendix *V* in the supplement.

Mechanism-based model

Our mechanism-based model focuses on a general tradeoff between competitive abilities of species and their abilities to access and retain nitrogen, adapted from a previous model suggested for Cedar Creek (Tilman 1994) to correspond to the monoculture traits we identified above. Species follow a strict competitive hierarchy, in which superior competitors pre-empt nitrogen uptake by inferior competitors, but inferior competitors do not affect superior competitors. Inferior competitors therefore persist only if they can access nitrogen that superior competitors cannot. Coexistence through this sort of resource partitioning is consistent with observations of increased plant cover and aboveground biomass in diverse plant communities at Cedar Creek (Fig. 1) (Tilman *et al.* 1997).

Competitive hierarchy is determined by species R^*, such that species j is competitively superior to species i if $R_j^* < R_i^*$. Annual nitrogen acquisition and retention is represented by species aboveground biomass, B_i, and nitrogen use efficiency, q_i. Arranging species from best competitor (i=1) to poorest, dynamics follow:

$$\frac{\mathrm{d}B_i}{\mathrm{dt}} = c_i B_i \left(1 - \frac{\sum_{j \le i} q_j B_j}{S}\right) - m_i B_i \qquad \text{Eq. (1)}$$

where c_i and m_i are species per-unit-biomass population growth and mortality rates, respectively, and S is total available nitrogen. Solving Eq. (1) for equilibrium monoculture biomass of species i, $B_i^*_{mono}$, yields:

$$B^*_{i,\mathrm{mono}} = \left(1 - \frac{m_i}{c_i}\right)\frac{S}{q_i} \qquad \text{Eq. (2)}$$

Eq. (2) shows that monoculture biomass in this model is determined by the ratio of population mortality and growth rates, which prevents species from obtaining greater than fraction $1–m_i/c_i$ of available nitrogen S.

Substituting Eq. (2) into Eq. (1) and solving for multi-species equilibrium biomass, B_i^*, shows that abundance of species i equals its monoculture abundance, less the biomass lost because of nitrogen acquired by superior competitors:

$$B_i^* = B^*_{i,\mathrm{mono}} - \sum_{j<i} \frac{q_j B_j^*}{q_i} \qquad \text{Eq. (3)}$$

Eq. (3) demonstrates that in order to persist in this model, poorer competitors must access a greater fraction of total available nitrogen than superior competitors do (i.e. species

with higher R^* must also have higher qB^*_{mono}; Fig. 2E). Eq. (3) therefore abstracts a number of potentially complex processes (m_i, c_i, and S) into a relationship among the three easily measured monoculture traits described above.

Stable coexistence

Because we assume a strict competitive hierarchy, Eq. (3) can be solved sequentially to determine equilibrium abundances for each species in a community, starting with the best competitor (i.e. i=1). Importantly, any equilibrium that is *feasible* (i.e. for which species have non-negative abundances) is also *globally stable*. In particular, if we assume that any species with equilibrium abundance less than zero is excluded from the mixture, then the equilibrium species abundances predicted by this model will always correspond to a single, analytically stable equilibrium. We demonstrate this property through linear stability analysis in Appendix *VI.i* in the supplement. A more rigorous proof of global stability is available in Takeuchi (1996 p. 36), Theorem 3.3.1, and (more briefly) in Takeuchi *et al.* (1978), Theorem 4.

Interestingly, our model's criteria for global stability are similar to those in MacArthur's 1970 resource competition model (MacArthur 1970), as utilized in Chesson (1990) to derive indicators of stable coexistence based on niche overlap (ρ) and fitness differences (f_2/f_1). Because of the strict competitive hierarchy in our model, we can also demonstrate its coexistence criteria using these two indicators (but see Haygood (2002) and Levine *et al*. (2017) for cases where pairwise frameworks are insufficient for identifying stable coexistence). Following Chesson (2000) and Letten *et al.* (2017), pairwise niche overlap and fitness differences in our model are defined as:

$$\rho = [(q_1/q_2)B_1^*{}_{mono}]^{1/2} \quad (4a)$$

$$f_2/f_1 = [(B_2^*{}_{mono})^2/((q_1/q_2)\, B_1^*{}_{mono})]^{1/2} \quad (4b)$$

Given that species 2 is the inferior competitor, stable coexistence occurs when $\rho<f_2/f_1$ (see Appendix *VI.ii* in the supplement for full derivation). Substituting Eqs. (4a-b) into this inequality and simplifying yields the criterion for stable coexistence:

$$(q_1/q_2)B_1^*{}_{mono} < B_2^*{}_{mono} \quad (5a)$$

Because inferior competitors have no effect on superior competitors in our model, we can substitute $q_1 B_1^*{}_{mono}$ in Eq. (5a) with the sum of nitrogen uptake by all superior competitors (i.e. species $j<i$), yielding the stability criteria for multi-species systems:

$$\Sigma_{j<i} (q_j/q_i) B_j^* < B_i^*{}_{mono} \quad (5b)$$

Note that Eq. (5b) is identical to the criterion for positive abundance (and thus stable coexistence) in Eq. (3).

Testing predictions

For each of the 122 experimental multi-species plots, we parameterized two versions of Eq. (3): one with raw traits, and one with traits snapped to the tradeoff surface. All models also included within-species trait variability, based on variation observed among replicate monocultures. We compared model outputs to observations from the multi-species mixtures based on predictions of stable coexistence, species-level abundance, and plot-level biomass. Detailed methods for making and testing model predictions are in Appendix *VII* and *VIII* in the supplement, respectively.

For each multi-species mixture, we used Eq. (5b) to determine which species were predicted to stably coexist. Observed coexistence was assumed for any species that appeared in at least one survey since 2001, which accounted for the small spatial scale of sampling relative to total plot area. Results from this method accord well with field observations in 2016, showing that almost all planted species persist in the multi-species plots, though often at low abundance.

For both species abundance and total plot biomass, we quantified mean absolute error (MAE), measured as |observation – prediction|, and report differences as "fold change". Each unit of fold change corresponds to a doubling or a halving – e.g. MAE = 1 implies that on average, predictions are either double or half the observed value. For species abundances, we calculated MAE for each plot as the mean error across all species, whereas for total plot-level aboveground biomass, MAE is the difference between observed and predicted productivity. For both plotting and analyses, we replaced predictions or observations of zero biomass with 0.02 g m^{-2}, corresponding to the lower 99.9% percentile of nonzero observations.

Extrapolating from the tradeoff surface

Lastly, we tested whether the tradeoff surface itself could be used to predict average relationships between community productivity and diversity. To conduct these tests, we constructed simulated "pools" of pseudo-species with traits for each simulated species determined by randomly choosing a point on the tradeoff surface. Because trait distributions differed between the 35 species grown in monoculture and the subset of those planted in the multi-species experiment, sampling was restricted to match the average trait distribution of functional groups that were present in the multi-species experiment (but not to match the trait distribution found in individual plots). We then predicted mean aboveground biomass and species abundance distributions for communities of 2-16 species drawn from these simulated species pools, and compared predictions to observations from the multi-species plots (details in Appendix *XI* in the supplement).

This test served two purposes. First, it tested whether our tradeoff surface and mechanism-based model successfully predicted positive relationships between diversity and productivity, which is a well-known property of grasslands at Cedar Creek (Tilman *et al.* 1997). Second, it demonstrated whether an empirically parameterized tradeoff surface could be used to make accurate predictions even if the identity and traits of individual species in mixtures are not known.

Results:

We found generally strong predictive power among the traits, tradeoffs, and mechanism-based models that we tested. Consistent with our hypotheses, species traits fell along a surface that matched tradeoffs required for multi-species coexistence in our theoretical model, and models parameterized with traits snapped to this surface substantially outperformed those parameterized with raw traits.

Regressions that used a single monoculture trait (such as just R*) generally provided poor fits for species abundance and aboveground biomass (Table 1). One exception was B^*_{mono}, which provided the best fits (R^2=0.20 for species abundances;

R^2=0.70 for aboveground biomass), though aboveground biomass fits were worse when planted richness was included as a covariate (R^2=0.58). Regression fits improved when all three monoculture traits were included, and improved substantially with two-way interactions among traits for both species abundance (R^2=0.48) and aboveground biomass (R^2=0.61). Three-way interactions did not significantly improve regression fits. In general, fits improved with planted richness as a covariate (two way interactions: R^2=0.63 for species abundances; R^2=0.64 for aboveground biomass).

Relationships among traits fell along a clear empirical tradeoff surface, closely matching a two-dimensional plane after transformation (Fig. 2A; R^2=0.93). Differences between raw and snapped trait values were small (Fig. 2B-D; $0.86 \leq R^2 \leq 0.97$). Species clustered in trait space by functional group, with high R^*, q, and B^*_{mono} for legumes, low values for C4 grasses, and intermediate values for C3 grasses and forbs. Species with "superior" values of any one trait had "inferior" value of one or both other traits (p<0.005; recall R^* is inversely related to competitive hierarchy). Moreover, interspecific trait variation was not well characterized by pairwise relationships (Fig. S2 in the supplement), suggesting that separate information about all three traits was necessary for characterizing tradeoffs.

Both mechanism-based models parameterized with raw and with snapped traits predicted stable coexistence among the vast majority of planted species (Fig. 3). Nevertheless, models parameterized with snapped traits predicted greater coexistence and better matched observations, particularly in diverse mixtures. In models parameterized with raw traits, prediction error was lower for total aboveground biomass than for species abundances, and error for abundances also increased at higher diversity (Fig. 4A,B). Nevertheless, models explained significant observed variation, and had relatively high goodness of fit for both abundances (Fig. 4C; R^2=0.23, p<0.001) and aboveground biomass (Fig. 4D; R^2=0.52, p<0.001). Models parameterized with snapped traits followed similar trends but with greater goodness of fit both for species abundances (Fig. 4E; R^2=0.46, p<0.001) and total aboveground biomass (Fig. 4F; R^2=0.59, p<0.001). For species abundances, snapped traits provided significantly better predictions across all richness treatments (Fig. 4A; p<0.01 for 4-species; p<0.001 for all others). For

aboveground biomass, predictions from snapped traits were significantly better for 8- and 16-species mixtures, and when considered on average across all diversity treatments (Fig. 4B; $p<0.05$ 8-species; $p<0.01$ 16-species; $p<0.001$ all treatments).

For simulated pools of pseudo-species with traits sampled randomly from the tradeoff surface, predictions closely matched observations for both aboveground biomass (Fig. 5A; $R^2=0.90$, $p<0.001$) and species abundance distributions (Fig. 5B-E; $0.52 \leq R^2 \leq 0.95$, $p<0.001$). Both observed and estimated aboveground biomass increased by roughly 30% for every doubling in planted diversity. Note that higher reported R^2 in Fig. 5 relative to Fig. 4 occurs because results compare means for each diversity treatment, rather than individual plots (rather than because of better fits per se). This is necessary because the simulated species pools do not correspond directly to any planted mixture.

Discussion:

Our results demonstrate that a combination of mechanism-based models and empirically observed tradeoffs can predict species abundances and productivity of multi-species communities. Because we do not tune model parameters to better fit predictions to observations, there is no a priori statistical reason to assume that snapping traits to an observed tradeoff surface should yield better model predictions. Our results therefore support the hypothesis that tradeoffs between species abilities to compete for nitrogen and to access and retain nitrogen – inherent in both our resource competition model and in the empirically observed relationships among species traits – are important in maintaining coexistence and determining species abundances at our site.

Tradeoffs and coexistence

The empirical tradeoff we identify separates functional groups in ways that are consistent with life histories. Based on monoculture traits, legumes are relatively poor nitrogen competitors (i.e. high R^*) but have high aboveground tissue nitrogen (i.e. high B^*_{mono} and q). This combination of traits likely results from nitrogen fixation, which allows legumes to access nitrogen that other species cannot, but also increases nitrogen availability for nearby species (Ranells & Wagger 1996). In contrast, traits identified C4

grasses as strong nitrogen competitors but poor at accessing nitrogen, consistent with experimental results suggesting that they competitively exclude many other species in low-nitrogen environments (Wedin & Tilman 1993; Dybzinski & Tilman 2007) but are less able to access nitrogen in deeper soil or during cooler growing seasons (McKane, Grigal & Russelle 1990). C3 grasses and forbs have intermediate trait values, though with slightly lower R^* and B^*_{mono} for C3 grasses, and a wider trait range for forbs. These intermediate traits suggest investment in strategies that increase access to additional nitrogen pools, but come at a cost to competitive ability. For example, many C3 grasses and forbs have fast vegetative growth rates and rapidly colonize new sites, which might allow them to exploit resources before superior competitors such as C4 grasses are able to reach them (Tilman 1994; Turnbull *et al.* 2013). Similarly, both C3 grasses and forbs tend to access deeper nitrogen pools and grow well in cooler seasons, reducing competitive interactions with C4 grasses (Fargione & Tilman 2005).

Coexistence in our model could result from many combinations of the strategies outlined above. The "coexistence mechanism" we identify is thus quite broad, encapsulating any strategies that allow species to access nitrogen that their competitors cannot. Nevertheless, our model provides some insight into how tradeoffs influence coexistence at Cedar Creek. Consider Eqs. (4a-b), which show that coexistence requires that fitness differences exceed niche overlap (Chesson 2000). Note that $(q_1/q_2)B_1^*_{mono}$ describes total nitrogen uptake by superior competitors (species 1) scaled by the tissue nitrogen concentration of inferior competitors (species 2), which is effectively the aboveground biomass of inferior competitors that is displaced by superior competitors. Eq. (4a) shows that niche overlap increases as $(q_1/q_2)B_1^*_{mono}$ grows, while Eq. (4b) shows that fitness differences decrease with $(q_1/q_2)B_1^*_{mono}$, or as monoculture biomass for the inferior competitor ($B_2^*_{mono}$) declines. These equations therefore imply that changes in traits that increase displacement of inferior competitors, such as increased nitrogen uptake by the superior competitor (i.e. higher q_1 or $B_1^*_{mono}$) or reduced tissue nitrogen concentrations in the inferior competitor (i.e. lower q_2), are likely to destabilize coexistence, whereas increased monoculture biomass of inferior competitors (i.e. higher $B_2^*_{mono}$) will tend to stabilize coexistence.

Though snapping traits to the tradeoff surface increased predictions of stable coexistence and more closely matched observations, these traits need not meet the criteria for coexistence in our model. While traits are negatively correlated on the surface, both the *shape* of the surface, and the *distribution* of traits across that surface can influence predictions of abundance and coexistence (McGill *et al.* 2006). Differences in the relative rates of change of q and B^*_{mono} along the fitted empirical tradeoff surface allow species to have lower total aboveground tissue nitrogen relative to superior competitors. We show this property mathematically in Appendix *X.i* in the supplement. Competitive exclusion is also demonstrated in Fig. 2E, where species do not all coexist, even with snapped traits.

Generalizations and broader applications

Though tradeoffs between nitrogen competition and retention explain substantial variation in coexistence, species abundances, and total community biomass at Cedar Creek, this result does not imply that these are the only mechanisms that structure plant community assembly. In Appendix *XI* in the supplement, we demonstrate how incorporating three specific strategies into our model results in significantly improved predictions. First, we incorporate experimental evidence which suggests that competitive hierarchies change in more fertile soils, potentially due to light competition (Dybzinski & Tilman 2007). Second, we explore different competitive hierarchies among legumes, which likely involve factors other than soil nitrogen (Ritchie & Tilman 1995). Lastly, we consider differences in phenology and rooting depth (McKane *et al.* 1990; Fargione & Tilman 2005), which likely add additional limits to the total amount of nitrogen uptake that "superior competitors" can pre-empt. Similarly, in other sites and systems, it is likely that entirely different traits and mechanisms will need to be considered to generate accurate predictions.

The fact that our model potentially omits important explanatory factors raises the question of whether it really constitutes a useful test of local coexistence mechanisms. We believe that it does, based on the improved predictive ability caused by snapping traits to the empirical tradeoff surface. The model parameterized with snapped traits performed almost as well as regressions including over a dozen fitted parameters. In the

fitted regression models, effect sizes and predictive power can be confounded by many well-known statistical artefacts, such as overfitting, covariance among variables, or influences of confounding variables (Adler *et al.* 2013; Kraft & Ackerly 2014). In contrast, the parameters in the mechanism-based model were not tuned to match predictions to observations, nor was the empirical tradeoff surface constrained to match theoretical relationships required for coexistence. It therefore appears highly unlikely that snapping traits to the tradeoff surface would improve predictive power without some link between hypothesized mechanisms, and those that actually drive community assembly.

Another advantage of the methods we apply is that tradeoffs are theoretically shared across all coexisting species (Tilman 1990, 2011; Chesson 2000), suggesting that tradeoff surfaces parameterized using subsets of species may generate accurate predictions for the entire community (McGill *et al.* 2006; Purves & Pacala 2008). Accurate predictions of species abundance distributions and community productivity from our simulated pools of pseudo-species support this hypothesis. Importantly, these simulations predicted higher productivity in multi-species mixtures than in any constituent monoculture (i.e. "transgressive over-yielding"), which is a well-known result from experiments at Cedar Creek (Tilman *et al.* 1997). Again, this is an emergent property – neither the tradeoff surface nor the model guaranteed this outcome (see Appendix *X.ii* in the supplement for details).

Caveats

There are several challenges limiting the applicability of the methods we present. First, many systems may be substantially more complex than the multi-species plots we consider, such that they cannot be approximated with simple mechanism-based models (Grimm, Ayllón & Railsback 2016). Similarly, available data may not always correspond to obvious empirical tradeoffs, nor to parameters in predictive mechanism-based models (Shipley *et al.* 2016). For example, though our fitted regressions show that including species richness as a covariate improved trait-based predictions of species abundances, it is unclear how we could empirically measure diversity dependent changes in monoculture traits such as R^*.

Though we may have been uniquely fortunate in being able to match long-term trait data to hypothesized coexistence mechanisms in our system, similar approaches have been successfully applied elsewhere to parameterize mechanism-based models (though without snapping traits), suggesting that there may be many systems in which our methods apply (Kneitel & Chase 2004; McGill *et al.* 2006; Pavlick *et al.* 2012; Litchman *et al.* 2012; Scheiter, Langan & Higgins 2013; Sakschewski *et al.* 2016). Moreover, new methods for analysing observational time series can identify causal associations among variables and estimate their effects on system dynamics through time (Sugihara *et al.* 2012; Deyle *et al.* 2016). These methods could be especially useful in systems where experimental data are lacking, or for parameterizing models for which only a subset of mechanisms can be linked to empirically observed traits and tradeoffs.

Conclusion

Our results show that information contained in empirically observed interspecific tradeoffs significantly improves predictions of coexistence, species abundances, and aboveground biomass in a perennial grassland. Though other locations or ecosystems may have different limiting factors and a different suite of relevant traits, an approach like ours that is based on empirical tradeoffs and mechanism-based models may help identify those factors and improve predictions. Because many physiological tradeoffs are broadly conserved across species and locations (Wright *et al.* 2004; Reich 2014; Kunstler *et al.* 2016), there may be similar overlap in the sets of traits and mechanisms that determine coexistence. If this is the case, then a manageably small number of trait-based measurements and models may yield strong predictions across a large range of spatial and temporal scales. The incorporation of information from tradeoffs into trait-based models of communities could therefore have the potential to make ecology an increasingly generalizable and predictive science.

Acknowledgments: This work was supported by grants from the US National Science Foundation Long-Term Ecological Research Program including DEB-8114302, DEB-8811884, DEB-9411972, DEB-0080382, DEB-0620652, and DEB-1234162. Further

support was provided by the Cedar Creek Ecosystem Science Reserve and the University of Minnesota. A.T.C. was supported by an NSF Graduate Research Fellowship, grant number 00006595, and by the Balzan Prize Foundation (awarded to D.T.). Computational resources were provided by the University of Minnesota Supercomputing Institute. Photographs in Fig. 1 were taken by J. Miller. We are grateful to S. Pacala for his suggestion to snap traits to a tradeoff surface to improve model fits. We also thank M. Burgess, M. Clark, J. Cowles, N. Eisenhauer, C. Farrior, G. Furey, B. Haegeman, F. Isbell, K. Kimmel, M. Loreau, C. Neuhauser, M. Thakur, and P. Wragg for advice and feedback on previous drafts of this manuscript, and S. Weisberg, for advice on statistical methods. We also thank several anonymous reviewers for their comments and suggestions, which helped us better clarify our methods, expand our interpretation of the mechanism-based model to encompass a wider range of strategies, and frame our discussions of coexistence in terms of Chesson's pairwise framework, which we hope make our results more broadly understandable.

Tables and Table Legends

				Partial R^2, linear model					*R^2_{adj}, Major Axis Regression*	
Model	**Type**	**n**	**p-value**	***B^*_{mono}***	***q***	***R****	**% soil C**	**Richness**	**abundance**	**total biomass**
null	base model	2	0.398				1		0.00	0.54
null	planted richness	4	< 0.001				< 0.001	0.037	0.10	0.46
B^*_{mono}	base model	4	< 0.001	0.038			< 0.001		0.10	**0.70**
B^*_{mono}	planted richness	8	< 0.001	0.038			0.001	0.036	0.20	0.58
q	base model	4	0.01		0.005		< 0.001		0.01	0.52
q	planted richness	8	0.054		0.004		0.001	0.036	0.10	0.33
*R**	base model	4	0.954			< 0.001	< 0.001		-0.01	0.54
*R**	planted richness	8	0.863			0.001	0.001	0.037	0.09	0.48
all traits	base model	8	< 0.001	0.085	0.001	0.017	0.001		0.26	**0.70**
all traits	planted richness	16	< 0.001	0.081	0.001	0.018	0.001	0.034	0.35	0.53
2-way interactions	base model	14	< 0.001	0.08	0.009	0.048	0.003		**0.48**	**0.61**
2-way interactions	planted richness	28	< 0.001	0.092	0.013	0.06	0.004	0.046	**0.63**	**0.64**
3-way interactions	base model	16	0.093	0.081	0.01	0.049	0.004		**0.49**	**0.60**
3-way interactions	planted richness	32	0.055	0.093	0.015	0.062	0.005	0.046	**0.63**	**0.65**

Table 1. Summary of linear regressions fitting species abundance in mixtures as a function of monoculture traits. All regressions include % soil carbon as a covariate. Regressions labelled with "planted richness" also include number of planted species as a covariate. Models labelled B^*_{mono}, q, or R^* include only a single trait, while others include all traits and their interactions. Column "n" shows number of parameters in the regression, p-values show results for ANOVA tests comparing nested models, and partial R^2 describes relative contribution of each term to total explanatory power. R^2_{adj} shows model fit from ranged major axis regression comparing observed and estimated values, either for species abundances, or for total community biomass for each planted mixture. Bolded results indicate regressions with R^2 greater than that for the mechanism-based model parameterized with snapped traits.

Figures and Figure Legends

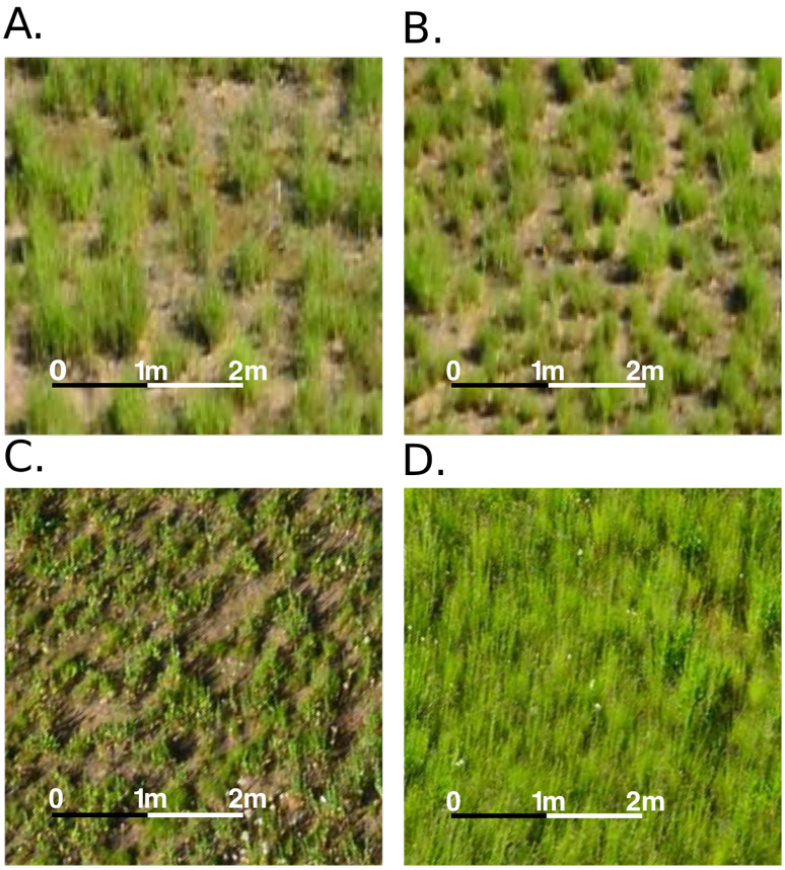

Figure 1: Relationships between vegetation density and species richness. In experimental multi-species mixtures of grassland plants at Cedar Creek, patchiness declines as planted diversity increases (**A-D** show 1, 2, 8, and 16 species respectively). This might suggest that some species persist by taking up resources that their competitors cannot access. (Photo credit J. Miller, 2014).

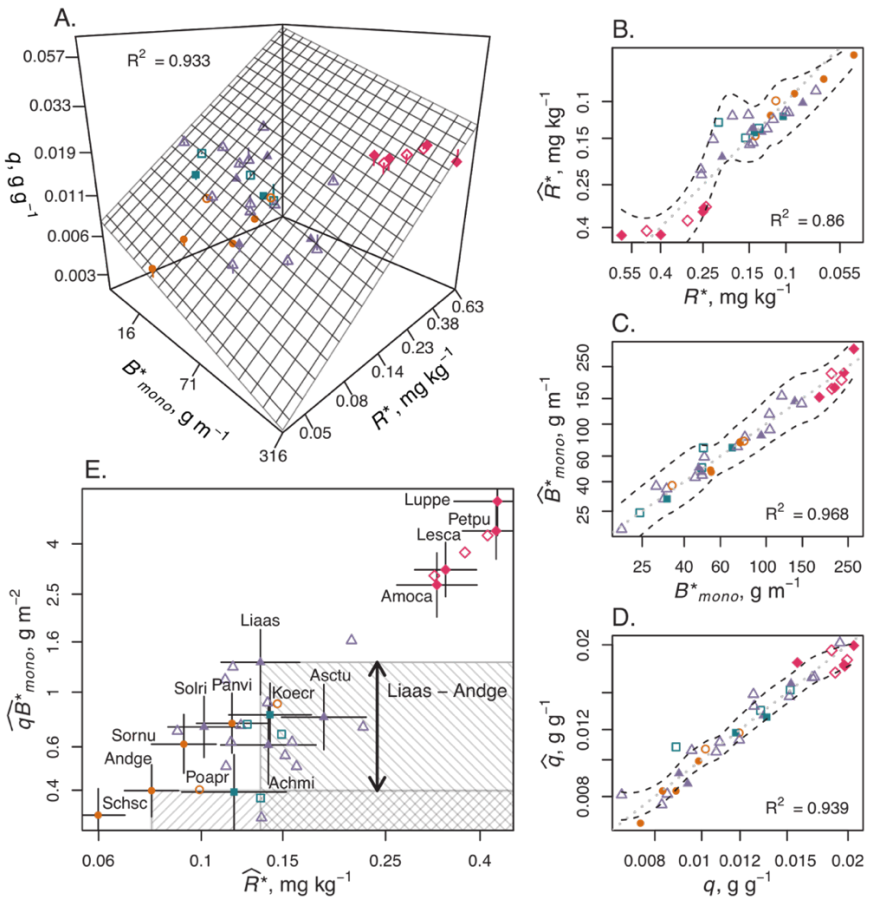

Figure 2: Tradeoffs for monoculture aboveground biomass B^*_{mono}, competitive hierarchy R^*, and proportion tissue nitrogen q. Traits snapped to the tradeoff surface are denoted by $\hat{B}^*_{mono}$, $\hat{R}^*$, and $\hat{q}$. Points show values for monocultures of 35 species. Green squares show C3 grasses, orange circles C4 grasses, purple triangles non-legume forbs, and pink diamonds legumes. Filled shapes denote species present in multi-species mixtures. **(A-D)** Raw traits fit closely to the tradeoff plane. **(E)** Species are excluded by competitors that have both lower R^* and higher qB^*_{mono}. Shaded regions show species excluded by *Andropogon gerardi* (Andge) or *Liatris aspera* (Liaas). Difference in qB^*_{mono} between species (e.g. "Liaas – Andge") shows the amount of nitrogen available to the inferior competitor. Dashed lines and intervals show mean ± one standard error for intraspecific trait variation.

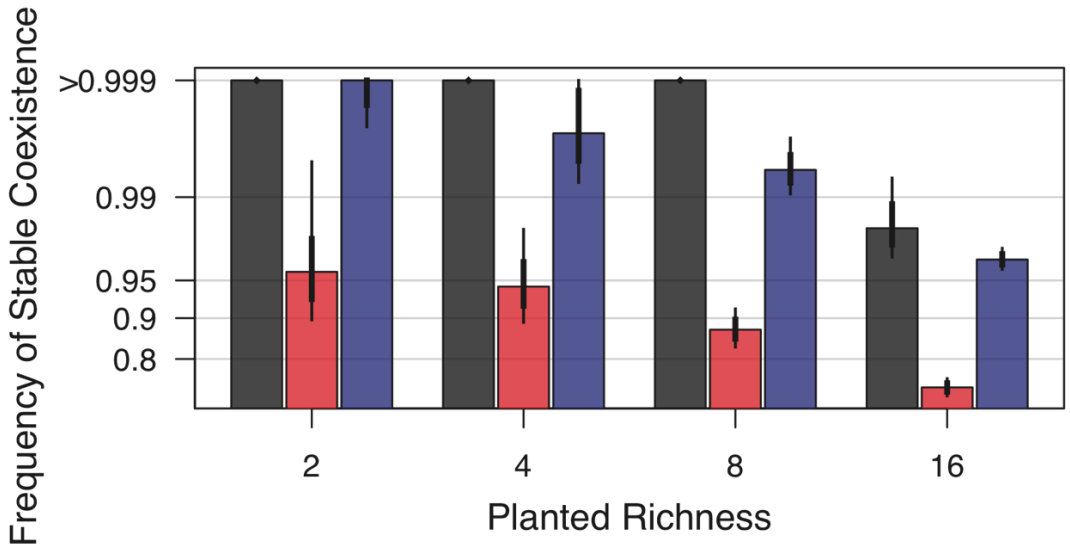

Figure 3: Coexistence in observed communities and model predictions. Bars and intervals show mean ± one standard deviation, and 95% confidence interval, averaged across observations or simulations. Black bars show fraction of planted species that have persisted in the multi-species plots. Red and blue bars show predictions of stable coexistence for mechanism-based models parameterized with raw and snapped traits, respectively. Stable coexistence in these models is identified following Eq. (5b), as described in the main text. See Fig. S3 in the supplement for species-level results.

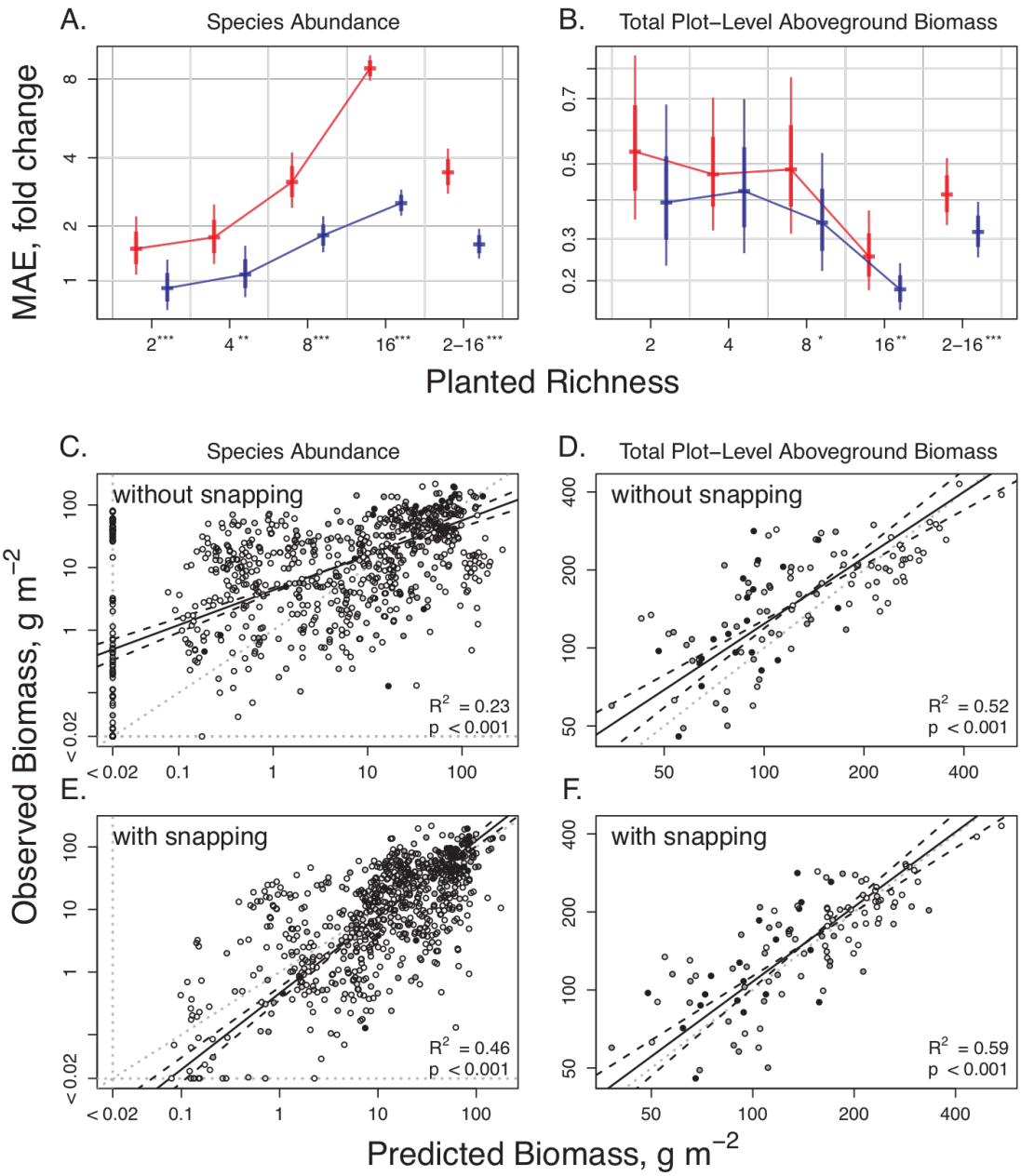

Figure 4: Performance of model predictions. **(A-B)** Mean absolute error (MAE) for predictions of species abundances and aboveground biomass across diversity treatment ("2-16" describes average fit across all treatments). Fold change indicates the number of doublings or halvings separating predictions and observations. Red and blue lines denote fits for raw and snapped traits, respectively. Intervals show 95% confidence interval and mean ± one standard deviation, averaged across simulations. Asterisks by labels signify significant differences between MAE from raw and snapped traits (*p<0.05; **p<0.01;

***p<0.001). **(C-F)** Comparison of observed and predicted species abundances and aboveground biomass for each multi-species plot. Dark to light-coloured shapes show predictions for 2, 4, 8, and 16-species mixtures, respectively. Dotted lines show lower detection limit ($0.02 gm^{-2}$) and 1-1 fit; solid and dashed lines show mean and 95% confidence interval for regression slope from ranged major axis regression. In **(A-B)**, p-values are from two-sample Wilcoxon tests; in **(C-F)**, p-values are from major axis regression.

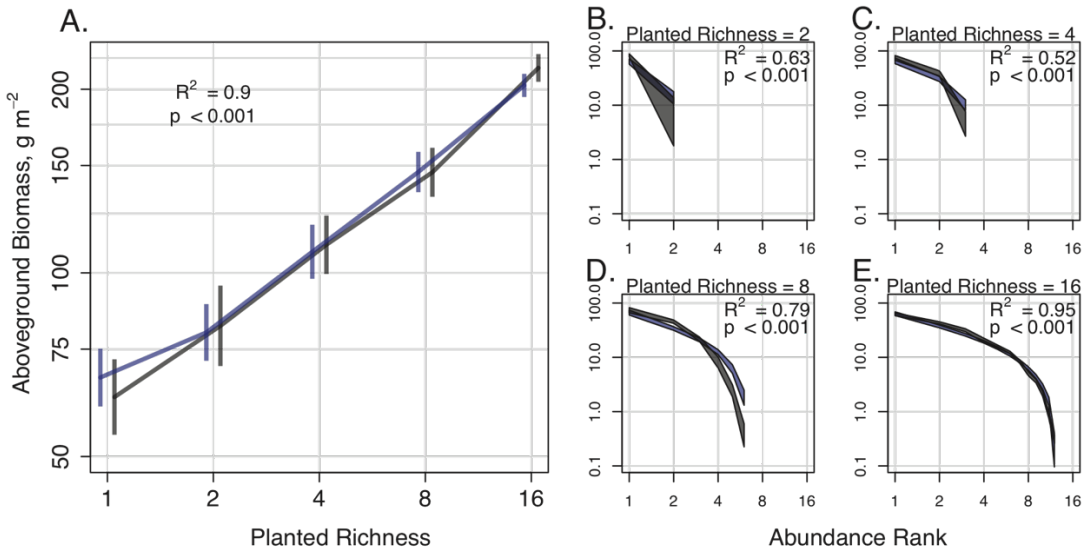

Figure 5: Results for simulated species pools generated by sampling traits by random draws of points from the empirical tradeoff surface. Black shows results from observed multi-species mixtures; blue shows simulation results. **(A)** Aboveground biomass increases with planted richness. Intervals show mean ± one standard error measured across plots. **(B-E)** Species abundance distributions across planted richness treatments. Shaded region shows mean ± one standard error measured across plots. "Abundance rank" sorts species from most to least abundant in each plot. R^2 and p-values are from major axis regression.

Bibliography:

Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013) Trait-based tests of coexistence mechanisms (ed C Scherber). *Ecology Letters*, **16**, 1294–1306.

Anderson, G.W., Guionnet, A. & Zeitouni, O. (2010) *An Introduction to Random Matrices*. Cambridge University Press, New York.

Armstrong, R.A. & McGehee, R. (1980) Competitive Exclusion. *The American Naturalist*, **115**, 151.

Asner, G.P., Elmore, A.J., Olander, L.P., Martin, R.E. & Harris, A.T. (2004) Grazing systems, ecosystem responses, and global change. *Annual Review of Environment and Resources*, **29**, 261–299.

Axelrod, A.N. & Irving, F.D. (1978) Some effects of prescribed fire at Cedar Creek Natural History Area. *Journal of the Minnesota Academy of Science*, **44**, 9–11.

Bartha, S., Meiners, S.J., Pickett, S.T.A. & Cadenasso, M.L. (2003) Plant colonization windows in a mesic old field succession. *Applied Vegetation Science*, **6**, 205–212.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**.

Bates, D. & Maechler, M. (2016) *Matrix: Sparse and Dense Matrix Classes and Methods.*

Bolker, B.M. & Pacala, S.W. (1999) Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *American Naturalist*, **153**, 575–602.

Borer, E.T., Seabloom, E.W., Gruner, D.S., Harpole, W.S., Hillebrand, H., Lind, E.M., Adler, P.B., Alberti, J., Anderson, T.M., Bakker, J.D., Biederman, L., Blumenthal, D., Brown, C.S., Brudvig, L.A., Buckley, Y.M., Cadotte, M., Chu, C., Cleland, E.E., Crawley, M.J., Daleo, P., Damschen, E.I., Davies, K.F., DeCrappeo, N.M., Du, G., Firn, J., Hautier, Y., Heckman, R.W., Hector, A., HilleRisLambers, J., Iribarne, O., Klein, J.A., Knops, J.M.H., La Pierre, K.J., Leakey, A.D.B., Li, W., MacDougall, A.S., McCulley, R.L., Melbourne, B.A., Mitchell, C.E., Moore, J.L., Mortensen, B., O'Halloran, L.R., Orrock, J.L., Pascual, J., Prober, S.M., Pyke, D.A., Risch, A.C., Schuetz, M., Smith, M.D., Stevens, C.J., Sullivan, L.L., Williams, R.J., Wragg, P.D., Wright, J.P. & Yang, L.H. (2014) Herbivores and nutrients control grassland plant diversity via light limitation. *Nature*, **508**, 517–520.

Bridle, J.R. & Vines, T.H. (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, **22**, 140–147.

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2004) Toward a Metabolic Theory of Ecology. *Ecology*, **85**, 1771–1789.

Burke, I.C., Lauenroth, W.K. & Coffin, D.P. (1995) Soil Organic Matter Recovery in Semiarid Grasslands: Implications for the Conservation Reserve Program. *Ecological Applications*, **5**, 793–801.

Carrara, F., Giometto, A., Seymour, M., Rinaldo, A. & Altermatt, F. (2015) Inferring species interactions in ecological communities: a comparison of methods at different levels of complexity (ed M Rees). *Methods in Ecology and Evolution*, **6**, 895–906.

Cavender-Bares, J., Gamon, J.A., Hobbie, S.E., Madritch, M.D., Meireles, J.E., Schweiger, A.K. & Townsend, P.A. (2017) Harnessing plant spectra to integrate the biodiversity sciences across biological and spatial scales. *American Journal of Botany*.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Cavender-Bares, J. & Reich, P.B. (2012) Shocks to the system: community assembly of the oak savanna in a 40-year fire frequency experiment. *Ecology*, **93**, S52–S69.

Chapin, F.S. (1980) The Mineral Nutrition of Wild Plants. *Annual Review of Ecology and Systematics*, **11**, 233–260.

Charpentier, A. (2015) Modelling Occurence of Events, with some Exposure. Blog Post. URL http://freakonometrics.hypotheses.org/20133 [accessed 6 June 2017]

Chase, J.M. & Leibold, M.A. (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago.

Chesson, P. (1990) MacArthur's consumer-resource model. *Theoretical Population Biology*, **37**, 26–38.

Chesson, P. (2000) Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology and Systematics*, **31**, 343–366.

Chesson, P. (2013) Species Competition and Predation. *Ecological Systems* (ed R. Leemans), pp. 223–256. Springer New York, New York, NY.

Clark, J.S., Bell, D., Chu, C.J., Courbaud, B., Dietze, M., Hersh, M., HilleRisLambers, J., Ibanez, I., LaDeau, S., McMahon, S., Metcalf, J., Mohan, J., Moran, E., Pangle, L., Pearson, S., Salk, C., Shen, Z.H., Valle, D. & Wyckoff, P. (2010) High-dimensional coexistence based on individual variation: a synthesis of evidence. *Ecological Monographs*, **80**, 569–608.

Clark, A.T., Rykken, J.J. & Farrell, B.D. (2011) The Effects of Biogeography on Ant Diversity and Activity on the Boston Harbor Islands, Massachusetts, USA. *Plos One*, **6**.

Clark, C.M. & Tilman, D. (2010) Recovery of plant diversity following N cessation: effects of recruitment, litter, and elevated N cycling. *Ecology*, **91**, 3620–3630.

Clements, F.E. (1916) *Plant Succession: An Analysis of the Development of Vegetation*. Carnegie Institution of Washington.

Connell, J.H. (1971) On the Role of Natural Enemies in Preventing Competitive Exclusion in Some Marine Animals and in Rain Forest Trees. *Dynamics of Populations* (eds P.J. Den Boer & G.R. Gradwell), p. Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands.

Cook, W.M., Yao, J., Foster, B.L., Holt, R.D. & Patrick, L.B. (2005) Secondary succession in an experimentally fragmented landscape: community patterns across space and time. *Ecology*, **86**, 1267–1279.

Cowles, J.M., Wragg, P.D., Wright, A.J., Powers, J.S. & Tilman, D. (2016) Shifting grassland plant community structure drives positive interactive effects of warming and diversity on aboveground net primary productivity. *Global Change Biology*, **22**, 741–749.

Coyte, K.Z., Schluter, J. & Foster, K.R. (2015) The ecology of the microbiome: Networks, competition, and stability. *Science*, **350**, 663–666.

Craine, J.M., Tilman, D., Wedin, D., Reich, P., Tjoelker, M. & Knops, J. (2002) Functional traits, productivity and effects on nitrogen cycling of 33 grassland species. *Functional Ecology*, **16**.

Cushing, E.J. (1963) *Late-Wisconsin Pollen Stratigraphy in East-Central Minnesota*. University of Minnesota, Minneapolis, MN, USA.

Davies, K.F., Margules, C.R. & Lawrence, J.F. (2004) A synergistic effect puts rare, specialized species at greater risk of extinction. *Ecology*, **85**, 265–271.

Deyle, E.R., May, R.M., Munch, S.B. & Sugihara, G. (2016) Tracking and forecasting ecosystem interactions in real time. *Proceedings of the Royal Society B: Biological Sciences*, **283**, 20152258.

Dormann, C.F. (2008) On community matrix theory in experimental plant ecology. *Web Ecology*, **8**, 108–115.

Dormann, C.F. & Roxburgh, S.H. (2005) Experimental evidence rejects pairwise modelling approach to coexistence in plant communities. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1279–1285.

Dutang, C. & Savicky, P. (2015) *Randtoolbox: Generating and Testing Random Numbers*.

Dybzinski, R. & Tilman, D. (2007) Resource use patterns predict long-term outcomes of plant competition for nutrients and light. *American Naturalist*, **170**.

Evans, M.E.K., Merow, C., Record, S., McMahon, S.M. & Enquist, B.J. (2016) Towards Process-based Range Modeling of Many Species. *Trends in Ecology & Evolution*, **31**, 860–871.

Fargione, J., Brown, C. & Tilman, D. (2003) Community assembly and invasion: An experimental test of neutral versus niche processes. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 8916–8920.

Fargione, J., Hill, J., Tilman, D., Polasky, S. & Hawthorne, P. (2008) Land Clearing and the Biofuel Carbon Debt. *Science*, **319**, 1235–1238.

Fargione, J. & Tilman, D. (2005) Niche differences in phenology and rooting depth promote coexistence with a dominant C-4 bunchgrass. *Oecologia*, **143**, 598–606.

Farrior, C.E., Tilman, D., Dybzinski, R., Reich, P.B., Levin, S.A. & Pacala, S.W. (2013) Resource limitation in a competitive context determines complex plant responses to experimental resource additions. *Ecology*, **94**, 2505–2517.

Fornara, D.A. & Tilman, D. (2008) Plant functional composition influences rates of soil carbon and nitrogen accumulation. *Journal of Ecology*, **96**, 314–322.

Foster, B.L. & Tilman, D. (2000) Dynamic and static views of succession: Testing the descriptive power of the chronosequence approach. *Plant Ecology*, 1–10.

Fukami, T., Martijn Bezemer, T., Mortimer, S.R. & Putten, W.H. (2005) Species divergence and trait convergence in experimental plant community assembly. *Ecology Letters*, **8**, 1283–1290.

Fukami, T. & Nakajima, M. (2011) Community assembly: alternative stable states or alternative transient states? *Ecology Letters*, **14**, 973–984.

Futuyma, D.J. (2010) Evolutionary constraint and ecological consequences. *Evolution*, **64**, 1865–1884.

Gause, G.F. (1932) Experimental studies on the struggle for existence I Mixed population of two species of yeast. *Journal of Experimental Biology*, **9**, 389–402.

Gause, G.F. (1934) Experimental analysis of Vito Volterra's Mathematical Theory of the Struggle for Existence. *Science*, **79**, 16–17.

Genz, A. & Bretz, F. (2009) *Computation of Multivariate Normal and T Probabilities*. Springer, Dordrecht ; New York.

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. & Hothorn, T. (2015) *Mvtnorm: Multivariate Normal and T Distributions. R Package Version 1.0-3.*

Gleason, H.A. (1926) The Individualistic Concept of the Plant Association. *Bulletin of the Torrey Botanical Club*, **53**, 7.

Gleeson, S.K. & Tilman, D. (1990) Allocation and the Transient Dynamics of Succession on Poor Soils. *Ecology*, **71**, 1144–1155.

Goh, B.S. (1977) Global Stability in Many-Species Systems. *The American Naturalist*, **111**, 135–143.

Gotelli, N.J. (2008) *A Primer of Ecology*, 4th ed. Sinauer Associates, Sunderland, Mass.

Gotelli, N.J., Anderson, M.J., Arita, H.T., Chao, A., Colwell, R.K., Connolly, S.R., Currie, D.J., Dunn, R.R., Graves, G.R., Green, J.L., Grytnes, J.-A., Jiang, Y.-H., Jetz, W., Kathleen Lyons, S., McCain, C.M., Magurran, A.E., Rahbek, C., Rangel, T.F.L.V.B., Soberón, J., Webb, C.O. & Willig, M.R. (2009) Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, **12**, 873–886.

Gould, S.J. & Lewontin, R.C. (1979) The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society B: Biological Sciences*, **205**, 581–598.

Grace, J.B., Anderson, T.M., Seabloom, E.W., Borer, E.T., Adler, P.B., Harpole, W.S., Hautier, Y., Hillebrand, H., Lind, E.M., Pärtel, M., Bakker, J.D., Buckley, Y.M., Crawley, M.J., Damschen, E.I., Davies, K.F., Fay, P.A., Firn, J., Gruner, D.S., Hector, A., Knops, J.M.H., MacDougall, A.S., Melbourne, B.A., Morgan, J.W., Orrock, J.L., Prober, S.M. & Smith, M.D. (2016) Integrative modelling reveals mechanisms linking productivity and plant species richness. *Nature*.

Grigal, D.W., Chamberlain, H.R., Finney, H.R., Wroblewski, D.V. & Gross, E.R. (1974) *Soils of the Cedar Creek Natural History Area*. University of Minnesota Agricultural Experiment Station, Saint Paul, MN, USA.

Grime, J.P. & Hunt, R. (1975) Relative Growth-Rate: Its Range and Adaptive Significance in a Local Flora. *The Journal of Ecology*, **63**, 393.

Grimm, V., Ayllón, D. & Railsback, S.F. (2016) Next-Generation Individual-Based Models Integrate Biodiversity and Ecosystems: Yes We Can, and Yes We Must. *Ecosystems*, **20**, 229–236.

Grimmett, G. & Stirzaker, D. (2001) *Probability and Random Processes*, 3rd ed. Oxford University Press, Oxford ; New York.

Haddad, N.M., Tilman, D. & Knops, J.M.H. (2002) Long-term oscillations in grassland productivity induced by drought. *Ecology Letters*, **5**, 110–120.

Harpole, W.S., Sullivan, L.L., Lind, E.M., Firn, J., Adler, P.B., Borer, E.T., Chase, J., Fay, P.A., Hautier, Y., Hillebrand, H., MacDougall, A.S., Seabloom, E.W., Williams, R., Bakker, J.D., Cadotte, M.W., Chaneton, E.J., Chu, C., Cleland, E.E., D'Antonio, C., Davies, K.F., Gruner, D.S., Hagenah, N., Kirkman, K., Knops, J.M.H., La Pierre, K.J., McCulley, R.L., Moore, J.L., Morgan, J.W., Prober, S.M., Risch, A.C., Schuetz, M., Stevens, C.J. & Wragg, P.D. (2016) Addition of multiple limiting resources reduces grassland diversity. *Nature*, **537**, 93–96.

Harpole, S. & Tilman, D. (2005) Non-neutral patterns of species abundance in grassland communities. *Ecology Letters*, **0**, 51017054245003.

Harpole, W.S. & Tilman, D. (2007) Grassland species loss resulting from reduced niche dimension. *Nature*, **446**, 791–793.

Harrison, R.D., Tan, S., Plotkin, J.B., Slik, F., Detto, M., Brenes, T., Itoh, A. & Davies, S.J. (2013) Consequences of defaunation for a tropical tree community (ed V Novotny). *Ecology Letters*, **16**, 687–694.

Hautier, Y., Niklaus, P.A. & Hector, A. (2009) Competition for Light Causes Plant Biodiversity Loss After Eutrophication. *Science*, **324**, 636–638.

Haygood, R. (2002) Coexistence in MacArthur-Style Consumer–Resource Models. *Theoretical Population Biology*, **61**, 215–223.

Hillebrand, H., Gruner, D.S., Borer, E.T., Bracken, M.E.S., Cleland, E.E., Elser, J.J., Harpole, W.S., Ngai, J.T., Seabloom, E.W., Shurin, J.B. & Smith, J.E. (2007) Consumer versus resource control of producer diversity depends on ecosystem type and producer community structure. *Proceedings of the National Academy of Sciences*, **104**, 10904–10909.

Hobbs, R. J. & Walker, L. R. (2007) Old Field Succession: Development of Concepts. *Old fields: dynamics and restoration of abandoned farmland*, The science and practice of ecological restoration (eds V.A. Cramer, R.J. Hobbs & Society for Ecological Restoration International), pp. 17–29. Island Press, Washington.

Hodapp, D., Hillebrand, H., Blasius, B. & Ryabov, A.B. (2016) Environmental and trait variability constrain community structure and the biodiversity-productivity relationship. *Ecology*, **97**, 1463–1474.

Huston, M. & Smith, T. (1987) Plant Succession: Life History and Competition. *The American Naturalist*, **130**, 168–198.

Hutchinson, G.E. (1959) Homage to Santa Rosalia or Why Are There So Many Kinds of Animals? *The American Society of Naturalists*, **93**, 145–159.

Inouye, R.S., Huntly, N.J., Tilman, D., Tester, J.R., Stillwell, M. & Zinnel, K.C. (1987) Old-Field Succession on a Minnesota Sand Plain. *Ecology*, **68**, 12–26.

Interlandi, S.J. & Kilham, S.S. (2001) Limiting resources and the regulation of diversity in phytoplankton communities. *Ecology*, **82**, 1270–1282.

Isbell, F., Tilman, D., Polasky, S., Binder, S. & Hawthorne, P. (2013) Low biodiversity state persists two decades after cessation of nutrient enrichment. *Ecology Letters*.

Janzen, D.H. (1970) Herbivores and the Number of Tree Species in Tropical Forests. *The American Naturalist*, **104**, 501.

Kitajima, K. & Tilman, D. (1996) Seed Banks and Seedling Establishment on an Experimental Productivity Gradient. *Oikos*, **76**, 381.

Kling, G., Hayhoe, K., Johnson, L., Magnuson, J., Polasky, S., Robinson, S., Shuter, B., Wander, M., Wuebbles, D. & Zak, D. (2003) *Confronting Climate Change in the Great Lakes Region: Impacts on Our Communities and Ecosystems*.

Kneitel, J.M. & Chase, J.M. (2004) Trade-offs in community ecology: linking spatial scales and species coexistence. *Ecology Letters*, **7**, 69–80.

Knops, J. & Tilman, D. (2000) Dynamics of soil nitrogen and carbon accumulation for 61 years after agricultural abandonment. *Ecology*, **81**, 88–98.

Knops, J.M.H., Tilman, D., Haddad, N.M., Naeem, S., Mitchell, C.E., Haarstad, J., Ritchie, M.E., Howe, K.M., Reich, P.B., Siemann, E. & Groth, J. (1999) Effects of plant species richness on invasion dynamics, disease outbreaks, insect abundances and diversity. *Ecology Letters*, **2**, 286–293.

Kraft, N.J.B. & Ackerly, D.D. (2014) Assembly of Plant Communities. *Ecology and the Environment* (ed R.K. Monson), pp. 67–88. Springer New York, New York, NY.

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. & Levine, J.M. (2015a) Community assembly, coexistence and the environmental filtering metaphor (ed J Fox). *Functional Ecology*, **29**, 592–599.

Kraft, N.J.B., Cornwell, W.K., Webb, C.O. & Ackerly, D.D. (2007) Trait Evolution, Community Assembly, and the Phylogenetic Structure of Ecological Communities. *The American Naturalist*, **170**, 271–283.

Kraft, N.J.B., Godoy, O. & Levine, J.M. (2015b) Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences*, **112**, 797–802.

Kunstler, G., Falster, D., Coomes, D.A., Hui, F., Kooyman, R.M., Laughlin, D.C., Poorter, L., Vanderwel, M., Vieilledent, G., Wright, S.J., Aiba, M., Baraloto, C., Caspersen, J., Cornelissen, J.H.C., Gourlet-Fleury, S., Hanewinkel, M., Herault, B., Kattge, J., Kurokawa, H., Onoda, Y., Peñuelas, J., Poorter, H., Uriarte, M., Richardson, S., Ruiz-Benito, P., Sun, I.-F., Ståhl, G., Swenson, N.G., Thompson, J., Westerlund, B., Wirth, C., Zavala, M.A., Zeng, H., Zimmerman, J.K., Zimmermann, N.E. & Westoby, M. (2016) Plant functional traits have globally consistent effects on competition. *Nature*, **529**, 204–207.

Lambers, H., Chapin, F.S. & Pons, T.L. (2008) *Plant Physiological Ecology*. Springer New York, New York, NY.

Lawson, D., Inouye, R.I., Huntley, N. & Carson, W.P. (1999) Patterns of woody plant abundance, recruitment, mortality, and growth in a 65 year chronosequence of old-fields. *Plant Ecology*, **145**, 267–279.

Lawton, J.H. (1999) Are There General Laws in Ecology? *Oikos*, **84**, 177.

Legendre, P. (2014) *lmodel2: Model II Regression*.

Lehman, C. (2000) *The Competition–Colonization Equations: Ecological and Evolutionary Implications*. PhD thesis, University of Minnesota, St. Paul, MN, USA.

Lehman, C., Loberg, S. & Clark, A.T. (2017) *Quantitative Ecology: A New Unified Approach*, 2713383118, 2017/03/28 ed. e-book.

Lehman, C.L. & Tilman, D. (2000) Biodiversity, Stability, and Productivity in Competitive Communities. *The American Naturalist*, **156**, 534–552.

Leskovec, J., Rajaraman, A. & Ullman, J.D. (2014) *Mining of Massive Datasets / Jure Leskovec, Standford University, Anand Rajaraman, Milliways Labs, Jeffrey David Ullman, Standford University*, Second edition. Cambridge University Press, Cambridge.

Letten, A.D., Ke, P.-J. & Fukami, T. (2017) Linking modern coexistence theory and contemporary niche theory. *Ecological Monographs*, **87**, 161–177.

Levin, S.A. (1970) Community Equilibria and Stability, and an Extension of the Competitive Exclusion Principle. *The American Naturalist*, **104**, 413–423.

Levine, J.M., Bascompte, J., Adler, P.B. & Allesina, S. (2017) Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, **546**, 56–64.

Li, S., Cadotte, M.W., Meiners, S.J., Hua, Z., Jiang, L. & Shu, W. (2015) Species colonisation, not competitive exclusion, drives community overdispersion over long-term succession (ed M Holyoak). *Ecology Letters*, **18**, 964–973.

Li, S., Cadotte, M.W., Meiners, S.J., Pu, Z., Fukami, T. & Jiang, L. (2016) Convergence and divergence in a long-term old-field succession: the importance of spatial scale and species abundance (ed M Rejmanek). *Ecology Letters*, **19**, 1101–1109.

Litchman, E., Edwards, K., Klausmeier, C. & Thomas, M. (2012) Phytoplankton niches, traits and eco-evolutionary responses to global environmental change. *Marine Ecology Progress Series*, **470**, 235–248.

Loreau, M. (2004) Does functional redundancy exist? *Oikos*, **104**, 606–611.

Lortie, C.J., Brooker, R.W., Choler, P., Kikvidze, Z., Michalet, R., Pugnaire, F.I. & Callaway, R.M. (2004) Rethinking plant community theory. *Oikos*, **107**, 433–438.

Lotka, A. (1932) Growth of mixed populations. *Journal of the Washington Academy of Science*, **22**, 461–469.

MacArthur, R.H. (1958) Population Ecology of Some Warblers of Northeastern Coniferous Forests. *Ecology*, **39**, 599–619.

MacArthur, R. (1970) Species packing and competitive equilibrium for many species. *Theoretical Population Biology*, **1**, 1–11.

MacArthur, R.H. (1972) *Geographical Ecology*. Harper & Row, New York.

MacArthur, R. & Levins, R. (1964) Competition, Habitat Selection, and Character Displacement in a Patchy Environment. *Proceedings of the National Academy of Sciences*, **51**, 1207–1210.

MacArthur, R.H. & Levins, R. (1967) Limiting Similarity Convergence and Divergence of Coexisting Species. *Am. Nat.*, **101**, 377–385.

Marsaglia, G. (2006) Ratios of Normal Variables. *Journal of Statistical Software*, **16**.

May, R.M. (1973) *Stability and Complexity in Model Ecosystems*, 1st Princeton landmarks in biology ed. Princeton University Press, Princeton.

May, R.M. & Anderson, R.M. (1987) Transmission dynamics of HIV infection. *Nature*, **326**, 137–142.

May, R.M. & MacArthur, R. (1972) Niche Overlap as a Function of Environmental Variability. *Proceedings of the National Academy of Sciences*, **69**, 1109–1113.

May, F., Wiegand, T., Lehmann, S. & Huth, A. (2016) Do abundance distributions and species aggregation correctly predict macroecological biodiversity patterns in tropical forests?: Species abundances, aggregation and biodiversity patterns. *Global Ecology and Biogeography*, **25**, 575–585.

McGill, B., Enquist, B., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, **21**, 178–185.

McKane, R.B., Grigal, D.F. & Russelle, M.P. (1990) Spatiotemporal differences in N-15 uptake and the organization of an old-field plant community. *Ecology*, **71**, 1126–1132.

Meiners, S.J., Pickett, S.T.A. & Cadenasso, M.L. (2015) *An Integrative Approach to Successional Dynamics: Tempo and Mode of Vegetation Change*. Cambridge University Press, Cambridge.

MHAPO. (2015) Minnesota Historical Aerial Photographs Online. URL https://www.lib.umn.edu/apps/mhapo/ [accessed 3 June 2017]

MHS. (2017) Minnesota Historical Society: Minnesota Treaty Interactive. URL http://usdakotawar.org/history/treaties/minnesota-treaty-interactive [accessed 8 June 2017]

Michalet, R., Chen, S., An, L., Wang, X., Wang, Y., Guo, P., Ding, C. & Xiao, S. (2015) Communities: are they groups of hidden interactions? (ed S Roxburgh). *Journal of Vegetation Science*, **26**, 207–218.

Miller, T.E. (1994) Direct and Indirect Species Interactions in an Early Old-Field Plant Community. *The American Naturalist*, **143**, 1007–1025.

Muller-Landau, H.C., Wright, S.J., Calderón, O., Condit, R. & Hubbell, S.P. (2008) Interspecific variation in primary seed dispersal in a tropical forest. *Journal of Ecology*, **96**, 653–667.

Murrell, D.J., Dieckmann, U. & Law, R. (2004) On moment closures for population dynamics in continuous space. *Journal of Theoretical Biology*, **229**, 421–432.

Myster, R.W. & Pickett, S.T.A. (1992) Dynamics of associations between plants in 10 old fields during 31 years of succession. *Journal of Ecology*, **80**, 291–302.

NADP. (2017) *National Atmospheric Deposition Program (NRSP-3)*. NADP Program Office, Illinois State Water Survey, University of Illinois, Champaign, IL, USA.

Naeem, S., Knops, J.M.H., Tilman, D., Howe, K.M., Kennedy, T. & Gale, S. (2000) Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. *Oikos*, **91**, 97–108.

Noble, I.R. & Slatyer, R.O. (1980) The use of vital attributes to predict successional changes in plant communities subject to recurrent disturbances. *Vegetatio*, **43**, 5–21.

Norden, N., Angarita, H.A., Bongers, F., Martínez-Ramos, M., Granzow-de la Cerda, I., van Breugel, M., Lebrija-Trejos, E., Meave, J.A., Vandermeer, J., Williamson, G.B., Finegan, B., Mesquita, R. & Chazdon, R.L. (2015) Successional dynamics in Neotropical forests are as uncertain as they are predictable. *Proceedings of the National Academy of Sciences*, **112**, 8013–8018.

Noy-Meir, I. (1975) Stability of Grazing Systems: An Application of Predator-Prey Graphs. *Journal of Ecology*, **63**, 459–481.

Palamara, G.M., Carrara, F., Smith, M.J. & Petchey, O.L. (2016) The effects of demographic stochasticity and parameter uncertainty on predicting the establishment of introduced species. *Ecology and Evolution*, **6**, 8440–8451.

Park, T. (1936) The effect of differentially conditioned flour upon the fecundity and fertility of Tribolium confusum duval. *Journal of Experimental Zoology*, **73**, 393–404.

Pavlick, R., Drewry, D.T., Bohn, K., Reu, B. & Kleidon, A. (2012) The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences Discussions*, **9**, 4627–4726.

Pickett, S.T.A. (1989) Space-for-Time Substitution as an Alternative to Long-Term Studies. *Long-Term Studies in Ecology* (ed G.E. Likens), pp. 110–135. Springer New York, New York, NY.

Pickett, S.T.A. & Cadenasso, M.L. (2005) Vegetation Dynamics. *Vegetation Ecology* (eds E. van der Maarel, J. Franklin & van der Maarel, Eddy), pp. 172–198. Blackwell Science Ltd, Oxford, UK.

Pickett, S.T.A., Cadenasso, M.L. & Bartha, S. (2001) Implications from the Buell-Small Succession Study for vegetation restoration. *Applied Vegetation Science*, **4**, 41–52.

Pierce, R.L. (1954) *Vegetation Cover Types and Land Use History of the Cedar Creek Natural History Reservation, Anoka and Isanti Counties, Minnesota*. University of Minnesota, Minneapolis, MN, USA.

Pimm, S.L. & Rosenzweig, M.L. (1981) Competitors and Habitat Use. *Oikos*, **37**, 1.

Post, W.M. & Kwon, K.C. (2000) Soil carbon sequestration and land-use change: processes and potential. *Global Change Biology*, **6**, 317–327.

Purschke, O., Schmid, B.C., Sykes, M.T., Poschlod, P., Michalski, S.G., Durka, W., Kühn, I., Winter, M. & Prentice, H.C. (2013) Contrasting changes in taxonomic, phylogenetic and functional diversity during a long-term succession: insights into assembly processes. *Journal of Ecol*.

Purves, D. & Pacala, S. (2008) Predictive Models of Forest Dynamics. *Science*, **320**, 1452–1453.

Pysek, P., Jarosik, V., Hulme, P.E., Pergl, J., Hejda, M., Schaffner, U. & Vila, M. (2012) A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. *Global Change Biology*, **18**.

R. Development Core Team,. (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Ramankutty, N. & Foley, J.A. (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997–1027.

Ranells, N.N. & Wagger, M.G. (1996) Nitrogen Release from Grass and Legume Cover Crop Monocultures and Bicultures. *Agronomy Journal*, **88**, 777.

Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Colwell, R.K. (2007) Species Richness and Evolutionary Niche Dynamics: A Spatial Pattern–Oriented Simulation Experiment. *The American Naturalist*, **170**, 602–616.

Reich, P.B. (2014) The world-wide "fast-slow" plant economics spectrum: a traits manifesto (ed H Cornelissen). *Journal of Ecology*, **102**, 275–301.

Reznick, D. (1985) Costs of Reproduction: An Evaluation of the Empirical Evidence. *Oikos*, **44**, 257.

Ricklefs, R.E. & Bermingham, E. (2002) The concept of the taxon cycle in biogeography. *Global Ecology and Biogeography*, **11**, 353–361.

Ritchie, M. & Tilman, D. (1995) Responses of legumes to herbivores and nutrients during succession on a nitrogen-poor soil. *Ecology*, **76**, 2648–2655.

Roff, D.A. (1992) Life History Theory: A Framework. *The Evolution of Life Histories*, pp. 35–61. Chapman & Hall.

Roff, D.A. & Fairbairn, D.J. (2007) The evolution of trade-offs: where are we? *Journal of Evolutionary Biology*, **20**, 433–447.

Roxburgh, S.H. & Wilson, J.B. (2000) Stability and coexistence in a lawn community: mathematical prediction of stability using a community matrix with parameters derived from competition experiments. *Oikos*, **88**, 395–408.

Rüger, N., Wirth, C., Wright, S.J. & Condit, R. (2012) Functional traits explain light and size response of growth rates in tropical tree species. *Ecology*, **93**, 2626–2636.

Sakschewski, B., von Bloh, W., Boit, A., Poorter, L., Peña-Claros, M., Heinke, J., Joshi, J. & Thonicke, K. (2016) Resilience of Amazon forests emerges from plant trait diversity. *Nature Climate Change*, **6**, 1032–1036.

Sanderson, M.A., Skinner, R.H., Barker, D.J., Edwards, G.R., Tracy, B.F. & Wedin, D.A. (2004) Plant species diversity and management of temperate forage and grazing land ecosystems. *Crop Science*, **44**, 1132–1144.

Schaffer, W.M. (1981) Ecological Abstraction: The Consequences of Reduced Dimensionality in Ecological Models. *Ecological Monographs*, **51**, 383–401.

Scheiter, S., Langan, L. & Higgins, S.I. (2013) Next-generation dynamic global vegetation models: learning from community ecology. *New Phytologist*, **198**, 957–969.

Schlaepfer, M.A., Runge, M.C. & Sherman, P.W. (2002) Ecological and evolutionary traps. *Trends in Ecology & Evolution*, **17**, 474–480.

Schoener, T.W. (1974) Some Methods for Calculating Competition Coefficients from Resource-Utilization Spectra. *The American Naturalist*, **108**, 332–340.

Shipley, B., De Bello, F., Cornelissen, J.H.C., Laliberté, E., Laughlin, D.C. & Reich, P.B. (2016) Reinforcing loose foundation stones in trait-based plant ecology. *Oecologia*, **180**, 923–931.

Soetaert, K., Petzoldt, T. & Setzer, R.W. (2010) Solving Differential Equations in R: Package deSolve. *Journal of Statistical Software*, **33**.

Stark, J.M. & Norton, J.M. (2015) The invasive annual cheatgrass increases nitrogen availability in 24-year-old replicated field plots. *Oecologia*, **177**, 799–809.

Staver, A.C., Archibald, S. & Levin, S.A. (2011) The Global Extent and Determinants of Savanna and Forest as Alternative Biome States. *Science*, **334**, 230–232.

Stearns, S.C. (1989) Trade-Offs in Life-History Evolution. *Functional Ecology*, **3**, 259.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M. & Munch, S. (2012) Detecting Causality in Complex Ecosystems. *Science*, **338**.

Takeuchi, Y. (1996) Global Stability. *Global dynamical properties of Lotka-Volterra systems*, pp. 21–56.

Takeuchi, Y., Adachi, N. & Tokumaru, H. (1978) Global stability of ecosystems of the generalized volterra type. *Mathematical Biosciences*, **42**, 119–136.

Tansley, A.G. (1935) The Use and Abuse of Vegetational Concepts and Terms. *Ecology*, **16**, 284–307.

Thomson, F.J., Moles, A.T., Auld, T.D. & Kingsford, R.T. (2011) Seed dispersal distance is more strongly correlated with plant height than with seed mass: Dispersal distance and seed mass. *Journal of Ecology*, **99**, 1299–1307.

Tilman, D. (1976) Ecological Competition Between Algae: Experimental Confirmation of Resource-Based Competition Theory. *Science*, **192**, 463–465.

Tilman, D. (1977) Resource Competition Between Planktonic Algae - Experimental and Theoretical Approach. *Ecology*, **58**, 338–348.

Tilman, D. (1982) *Resource Competition and Community Structure*. Princeton University Press, Princeton, NJ.

Tilman, D. (1987) Secondary succession and the pattern of plant dominance along experimental nitrogen gradients. *Ecological Monographs*, **57**, 189–214.

Tilman, D. (1990) Constraints and Tradeoffs - toward a Predictive Theory of Competition and Succession. *Oikos*, **58**, 3–15.

Tilman, D. (1994) Competition and biodiversity in spatially structured habitats. *Ecology*, **75**, 2–16.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10854–10861.

Tilman, D. (2011) Diversification, Biotic Interchange, and the Universal Trade-Off Hypothesis. *American Naturalist*, **178**, 355–371.

Tilman, D. & Haddi, A.E. (1992) Drought and diversity in grasslands. *Oecologia*, **89**, 257–264.

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M. & Siemann, E. (1997) The influence of functional diversity and composition on ecosystem processes. *Science*, **277**, 1300–1302.

Tilman, D., Reich, P.B. & Isbell, F. (2012) Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory. *Proceedings of the National Academy of Sciences*, **109**, 10394–10397.

Tilman, D., Wedin, D. & Knops, J. (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, **379**, 718–720.

Tjur, T. (2009) Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination. *The American Statistician*, **63**, 366–372.

Turnbull, L.A., Levine, J.M., Loreau, M. & Hector, A. (2013) Coexistence, niches and biodiversity effects on ecosystem functioning (ed JHR Lambers). *Ecology Letters*, **16**, 116–127.

Uhl, C., Buschbacher, R. & Serrao, E.A.S. (1988) Abandoned Pastures in Eastern Amazonia. I. Patterns of Plant Succession. *The Journal of Ecology*, **76**, 663.

USDA. (2017) USDA PLANTS Database. URL https://plants.usda.gov/java/ [accessed 8 June 2017]

Vandermeer, J.H. (1969) The Competitive Structure of Communities: An Experimental Approach with Protozoa. *Ecology*, **50**, 362–371.

Vellend, M. (2016) *The Theory of Ecological Communities*. Princeton University Press, Princeton.

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.

Volterra, V. (1926) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Mem. Acad. Lincei Roma.*, **2**, 31–113.

van de Voorde, T.F.J. van de, van der Putten, W.H. & Bezemer, T.M. (2011) Intra- and interspecific plant-soil interactions, soil legacies and priority effects during old-field succession: Intra- and interspecific plant-soil interactions. *Journal of Ecology*, **99**, 945–953.

Walker, L.R., Wardle, D.A., Bardgett, R.D. & Clarkson, B.D. (2010) The use of chronosequences in studies of ecological succession and soil development: Chronosequences, succession and soil development. *Journal of Ecology*, **98**, 725–736.

Wangersky, P.J. (1978) Lotka-Volterra Population Models. *Annual Review of Ecology and Systematics*, **9**, 189–218.

Webb, L.J., Tracey, J.G. & Williams, W.T. (1972) Regeneration and Pattern in the Subtropical Rain Forest. *The Journal of Ecology*, **60**, 675.

Wedin, D. & Tilman, D. (1993) Competition Among Grasses Along a Nitrogen Gradient: Initial Conditions and Mechanisms of Competition. *Ecological Monographs*, **63**, 199.

Weigelt, A., Schumacher, J., Walther, T., Bartelheimer, M., Steinlein, T. & Beyschlag, W. (2007) Identifying mechanisms of competition in multi-species communities. *Journal of Ecology*, **95**, 53–64.

Whisenant, S.G., Thurow, T.L. & Maranz, S.J. (1995) Initiating Autogenic Restoration on Shallow Semiarid Sites. *Restoration Ecology*, **3**, 61–67.

Whittington, H.R., Tilman, D. & Powers, J.S. (2013) Consequences of elevated temperatures on legume biomass and nitrogen cycling in a field warming and biodiversity experiment in a North American prairie. *Functional Plant Biology*, **40**, 1147.

Wilbur, H.M. (1972) Competition, Predation, and the Structure of the Ambystoma-Rana Sylvatica Community. *Ecology*, **53**, 3–21.

Wilson, E.O. (1961) The Nature of the Taxon Cycle in the Melanesian Ant Fauna. *The American Naturalist*, **95**, 169–193.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models: Estimation of Semiparametric Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**, 3–36.

Wright, J.S. (2002) Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia*, **130**, 1–14.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.

Wright, S.J., Trakhtenbrot, A., Bohrer, G., Detto, M., Katul, G.G., Horvitz, N., Muller-Landau, H.C., Jones, F.A. & Nathan, R. (2008) Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. *Proceedings of the National Academy of Sciences*, **105**, 19084–19089.

Zanini, L., Ganade, G. & Hübel, I. (2006) Facilitation and competition influence succession in a subtropical old field. *Plant Ecology*, **185**, 179–190.

Zera, A.J. & Harshman, L.G. (2001) The Physiology of Life History Trade-Offs in Animals. *Annual Review of Ecology and Systematics*, **32**, 95–126.

Zumkehr, A. & Campbell, J.E. (2013) Historical U.S. Cropland Areas and the Potential for Bioenergy Production on Abandoned Croplands. *Environmental Science & Technology*, **47**, 3840–3847.

Appendices for Chapter One:

Determinism and stochasticity during 88 years of grassland succession:
Roles of soil fertility, fire, climate, competition, dispersal, and mortality

Appendix A: Site and species data

Data from the plant abundance and percent-cover old field surveys is included in the supplement to this paper. All data, including environmental covariates, can also be accessed through the Cedar Creek data webpage (cedarcreek.umn.edu/research/data), or the LTER data portal (portal.lternet.edu). Full methods, diagrams of plot layouts, and lists of specific burning years and season for each field can be found on the Cedar Creek webpage, under Experiment 014 (cedarcreek.umn.edu/research/data/methods?e014) and Experiment 054 (cedarcreek.umn.edu/research/data/methods?e054).

A.1: Study design and sampling

Two types of surveys of vegetative dynamics are conducted in the old fields at Cedar Creek. Percent-cover surveys ("Experiment 014") have been conducted eight times since 1983, at roughly five year intervals. Specific survey years are 1983, 1989, 1994, 1997, 2002, 2006, 2011, and 2016. Aboveground biomass data has been collected annually since 1988 ("Experiment 054"). Sampling methods are descried in the main text, and in Inouye et al. (1987) (Experiment 14), Tilman (1987) (Experiment 14), and in Haddad et al. (2002) (Experiment 54). Soil total nitrogen concentration has been measured in the percent-cover plots every six years since the start of the experiment, and has been collected twice in the biomass plots (2001 and 2014). Methods are described in the main text, and in Knops and Tilman (2000). Vegetation and soil surveys are ongoing as part of long-term research at Cedar Creek.

A.2: Species data

To help reduce bias from changes in surveyor identity over time, we produced a simplified species list, lumping together species that are commonly mistaken for one another, or that are otherwise difficult to distinguish. We then grouped these by functional group, duration, and provenance following the methods described in the main

text. In general, we used the same groupings described by Inouye et al. (1987). Where there was disagreement among sources, or where data were otherwise unavailable, we used records for Minnesota listed in the USDA plants database (USDA 2017).

A.3: Nitrogen data

Soil nitrogen concentrations have strong temporal dynamics in old fields and Cedar Creek, and generally increase with field age (Knops & Tilman 2000). Because soil and vegetation surveys were not always conducted in the same years, we interpolated observed soil nitrogen dynamics in order to approximate concentrations at the time of vegetation sampling. For percent-cover plots, where we had multiple observations per plot, we used the `lmer` function in the `lme4` package (version 1.1-13) in `R` (version 3.3.2) (Bates *et al.* 2015; R. Development Core Team, 2016) to fit a model of the form:

```
lmer(log(soilN)~log(age)+(1|field/transect/plot))
```

where `soilN` is percent soil total nitrogen (%N in oven-dried soil, on a mass by mass basis), `age` is years since field abandonment, and `field`, `transect`, and `plot` signify the sampling units of each field.

For the biomass plots, where we had only two measurements, we fit a linearized model of soil nitrogen dynamics following the methods of Knops and Tilman (2000). Specifically, we used the equation:

$r_i = \log(soilN_{14,i}/soilN_{01,i})$

where r_i is the exponential growth rate of soil nitrogen concentrations in plot i, $soilN_{01,i}$ is the soil nitrogen concentration in 2001, and $soilN_{14,i}$ is the concentration in 2014. We then calculated soil nitrogen concentration in plot i at time t, where t is the calendar year, as:

$soilN_{t,i} = soilN_{01,i} \exp(r_i [t - 2001])$

For both types of models, we filled in missing data for plots based on the lowest nesting level possible (i.e. transect, field, or global mean).

Appendix B: Regression methods

B.1: Interpreting model coefficients

For both regression models, we present standardized coefficients. These describe the statistical effect of changing each variable by a unit of one standard deviation (e.g. increasing soil nitrogen concentration from the mean level, to the mean level plus one standard deviation). To make changes in the response variable more interpretable, we present coefficients for the biomass model scaled in terms of fold change (i.e. the fraction change in biomass resulting from a change of one standard deviation in one of the explanatory variables, while holding all others constant). For the colonization and mortality models, we present coefficients in terms of hazard ratios, as described in the main text.

As noted in the main text, a major advantage of the complementary log-log link function is that an offset term in the model can be used to account for unequal time spans between survey intervals. This is shown nicely in a proof by Charpentier (2015), which we reproduce here. In the complementary log-log link function, a linear model is transformed into probability units as:

$p = 1 - \exp(-\exp(y))$

where p is the probability of an event, and y is the untransformed output of the linear model. Now, imagine that we wish to calculate an annual probability p (e.g. annual probability of mortality) given some observation of n years. If p_n is the probability of a species surviving for n years, then we can relate this to p as:

$p_n = 1 - (1 - p)^n$

In other words, one minus the probability of a species not dying or not colonizing a plot for n consecutive years. Note that we can re-write this as:

$p_n = 1 - \exp(\log(1 - p)\ n)$

Thus, we can substitute this into the link function as:

$p_n = 1 - \exp(\log(1 - [1 - \exp(-\exp(y))])\ n)$

$p_n = 1 - \exp(\log(-\exp(-\exp(y)))\ n)$

$p_n = 1 - \exp(-\exp(y + \log(n)))$

This shows that if we include $\log(n)$ as an offset (n.b. an offset is simply a constant term added to a regression), we can account for differential sampling intervals (often called

"exposure times") in our binomial model. This is the method we use for the regressions presented in the main text.

B.2: Stepwise selection methods

To simplify our regressions, we used a stepwise selection procedure. For both types of regressions, we started with the full model described in the main text. For the binomial models, we then removed terms one at a time, and kept the model with the lowest AIC. We then repeated the procedure until removing terms no longer decreased AIC. For the biomass model, we used a similar approach, but with adjusted-R^2 (which accounts for number of parameters in the model) rather than AIC, as this was simpler to calculate for the GAMMs that we used.

B.3: Model structure

Specific model forms resulting from the stepwise selection procedure are listed below. For the abundance regressions, using the `gamm` function `mgcv` package (version 1.8-17) in `R` (Wood 2011), we arrived at the following models:

```
C4s~s(age,by=field)+ soilN + burning + precip*temp + C3s
C3s~s(age,by=field)+ soilN + precip*temp + C4s
Annuals~s(age,by=field)+ burning + precip*temp + C4s + C3s
Forbs~s(age,by=field)+ soilN + burning + precip + temp +
C3s
Legumes~s(age)+ precip*temp + C4s
```

Model coefficients and species groups are as described in the main text. All models also included the random effects term:

```
random=list(field=~1,transect=~1)
```

where `field` and `transect` describe the sampling units. Note that the legume model did not include separate splines for each field, as this model did not converge.

For colonization and mortality regressions, we used the `glmer` function in the `lme4` package. For all of these models, we used a complementary log-log link function, using the command:

`family = binomial(link="cloglog")`

We also included an offset term of the form:

`offset(log(dt_m1))`

and included a random effect of the form:

`(1|field)`

For colonization models, we arrived at the forms:

```
C4s~soilN + burning + precip*temp + C3s + trnsabund
C3s~precip + C4s + trnsabund
Annuals~soilN + burning + precip*temp + C4s + C3s +
trnsabund
1 – Forbs~soilN + burning + precip*temp + C3s + trnsabund
Legumes~soilN + burning + precip*temp + C4s + trnsabund
```

where "`trnsabund`" is the within-group mean abundance in each transect. Note that for forbs, we had to fit the model for 1 – probability of colonization, as probability of colonization would not converge. This is a disadvantage of the complementary log-log link function, as some ranges of parameter values can lead to probability estimates outside of the range [0,1], and therefore cause model convergence failure. Though this meant that we could not include an offset for the forb model, we nevertheless transformed predictions to provide estimates of annual colonization probability, based on mean inter-survey time period.

For mortality models, we arrived at the forms:

```
C4s~soilN + burning + precip*temp + C3s + pltabund
C3s~burning + precip*temp + C4s + pltabund
Annuals~soilN + burning + precip*temp + C4s + C3s +
pltabund
Forbs~soilN + burning + precip*temp + C4s + C3s + pltabund
Legumes~soilN + burning + precip*temp + C4s + pltabund
```

where "`pltabund`" is the within-group abundance in each plot during the previous survey.

For the abundance model, we used the default adjusted R^2 estimate generated by the `gamm` function. For the binomial regressions, we used the pseudo-R^2 index proposed

by Tjur (2009), which is simply the difference between the mean predicted probability of an event occurring for all instances where the event was observed to occur, minus the mean predicted probability for all events where the event was not observed to occur. Note that when predictions perfectly identify events, this index is 1, and decreases towards a minimum value of zero as predictions get worse.

Appendix C: Metapopulation simulation

C.1: Model methods

As described in the main text, we conducted the metapopulation simulations using predictions from the abundance, colonization, and mortality regressions. For each time step, we calculated predicted abundance of each group, then simulated stochastic mortality events, and then simulated stochastic colonization events (based on the pre-mortality abundance). At year zero (i.e. at the beginning of the simulation), we assumed zero abundance of all species.

Specific details about the scenarios are described in the main text. Note that to generate the empirical relationship between precipitation and temperature at Cedar Creek, we fit a ranged major axis regression using the `lmodel2` function (version 1.7-2) in `R` (Legendre 2014). This resulted in the relationship:

```
precip = 1963.772 - 63.99626*temp
```

where `precip` is average total precipitation in mm, and `temp` is mean daily high in degrees C, measured from April to August.

Each simulation run translates roughly to mean dynamics at the level of a percent-cover survey transect (i.e. 25 simulated "plots"). For each scenario, we simulated 1,000 iterations. In general, this process was very simple, and required nothing more than careful book keeping of how the simulation responded to predictions from the regressions. One exception to this was the interactive effects of C3 and C4 competition. To account for their negative effects on one another's abundances, we jointly solved the linear equations predicting each of their abundances (i.e. assuring that C3s and C4s were predicted to have the same abundance as was used to calculate their negative effects on

other groups). In cases where we predicted negative abundances for either group, we set their abundance to zero.

C.2: Full scenario results

In addition to the scenarios described in the main text, we simulated a set of fully interactive scenarios of burning, soil fertility, and climate in order to show how predictions from our metapopulation model changed across a wider range of conditions. For burning, we simulated two scenarios: one that had been burned regularly since field abandonment, and one that was never burned. For climate, we simulated five scenarios with mean climate conditions drawn from the major axis regression line comparing historical precipitation and temperature trends. These corresponded to mean temperature conditions falling between zero and two standard deviations above and below the historical mean, and inter-annual variability determined by the observed historical variance in both climate variables, roughly spanning the most extreme events observed in the past 50 years at Cedar Creek (including the 1988 "North American Drought" and the 1993 "Great Mississippi Flood"; see Fig. 2b in the main text). Lastly, we used our fitted regression of soil total nitrogen dynamics for the percent-cover plots to generate three scenarios of nitrogen dynamics: one following the mean relationship of increasing soil total nitrogen concentrations with field age, and two following the mean trend ± one standard deviation, based on observed among-field variability (Fig. 2a in the main text). Unlike the scenarios described in the main text, we did not vary soil nitrogen dynamics among simulations within individual scenarios for these expanded analyses.

Because plotting observed successional trajectories for each of the 30 interactive combinations of scenarios outlined above (two burning, three soil fertility, and five climate), we used four indices to summarize successional dynamics (see examples in Fig. S5). $\mathrm{Annual}_{\mathrm{mid.}}$ shows final year for which annual forbs and grasses make up at least 50% of total aboveground biomass; $\mathrm{C4}_{\mathrm{mid.}}$ shows first year for which C4 grasses and sedges make up at least 50% of biomass; $\mathrm{C3}_{\mathrm{range}}$ shows range of years for which C3 grasses make up at least 50% of biomass; and e^{H}_{mid} shows mean Shannon diversity of species groups averaged across all years (Fig. S5). These represent, respectively, the length of

time that annuals persist before they are displaced by other species, the length of time required for C4s to colonize and dominate a field, the length of time that C3s exclude other species from a field, and the general abundance of groups of rarer, non-dominant species groups (i.e. forbs and legumes).

Results largely accorded with those from the simpler set of scenarios described in the main text (Fig. S6). Depending on the scenario, Annual_{mid} varied from 5-15 years, $\text{C4}_{mid.}$ from 20-70 years, C3_{range} from 15-60 years, and e^{H}_{mid} from 1.9-2.8. Warmer, dryer conditions tended to increase dominance time by annuals (i.e. larger Annual_{mid}), decrease the time it took C4s to become dominant (i.e. smaller $\text{C4}_{mid.}$), and decrease the length of time that C3s were dominant (i.e. smaller C3_{range}). Mean species group diversity (i.e. e^{H}_{mid}) followed a humped-shaped relationship, with highest diversity given current mean climate conditions. Lastly, both burning and reduced soil total nitrogen concentrations had similar impacts, with faster dominance by C4s, decreased dominance time by C3s, increased diversity, and relatively little effect on annuals.

Figures:

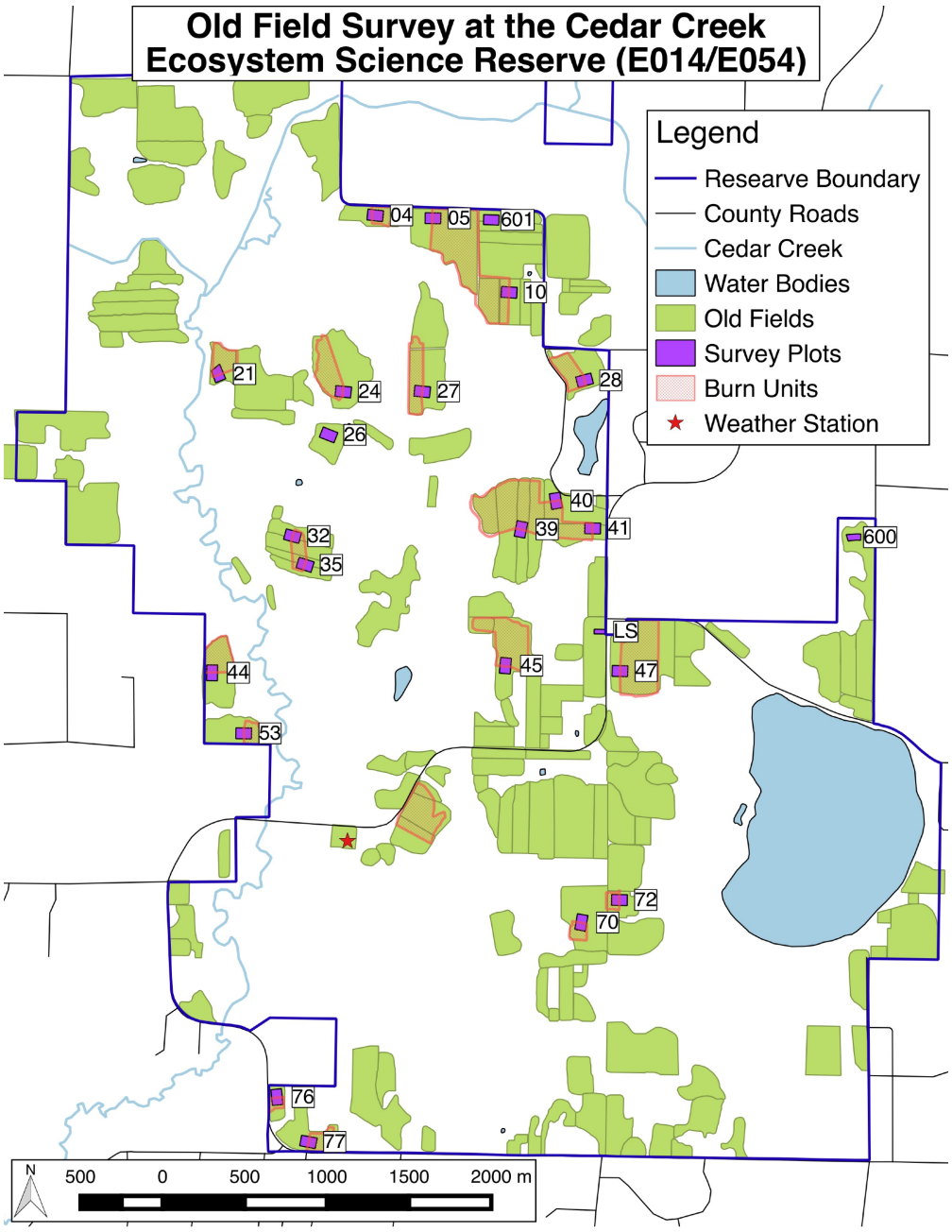

Figure S1: Map of the old field experiments at the Cedar Creek Ecosystem Science Reserve. Purple rectangles show regions with survey plots in each field. Numbers by each rectangle correspond to field identifiers in Table 1.

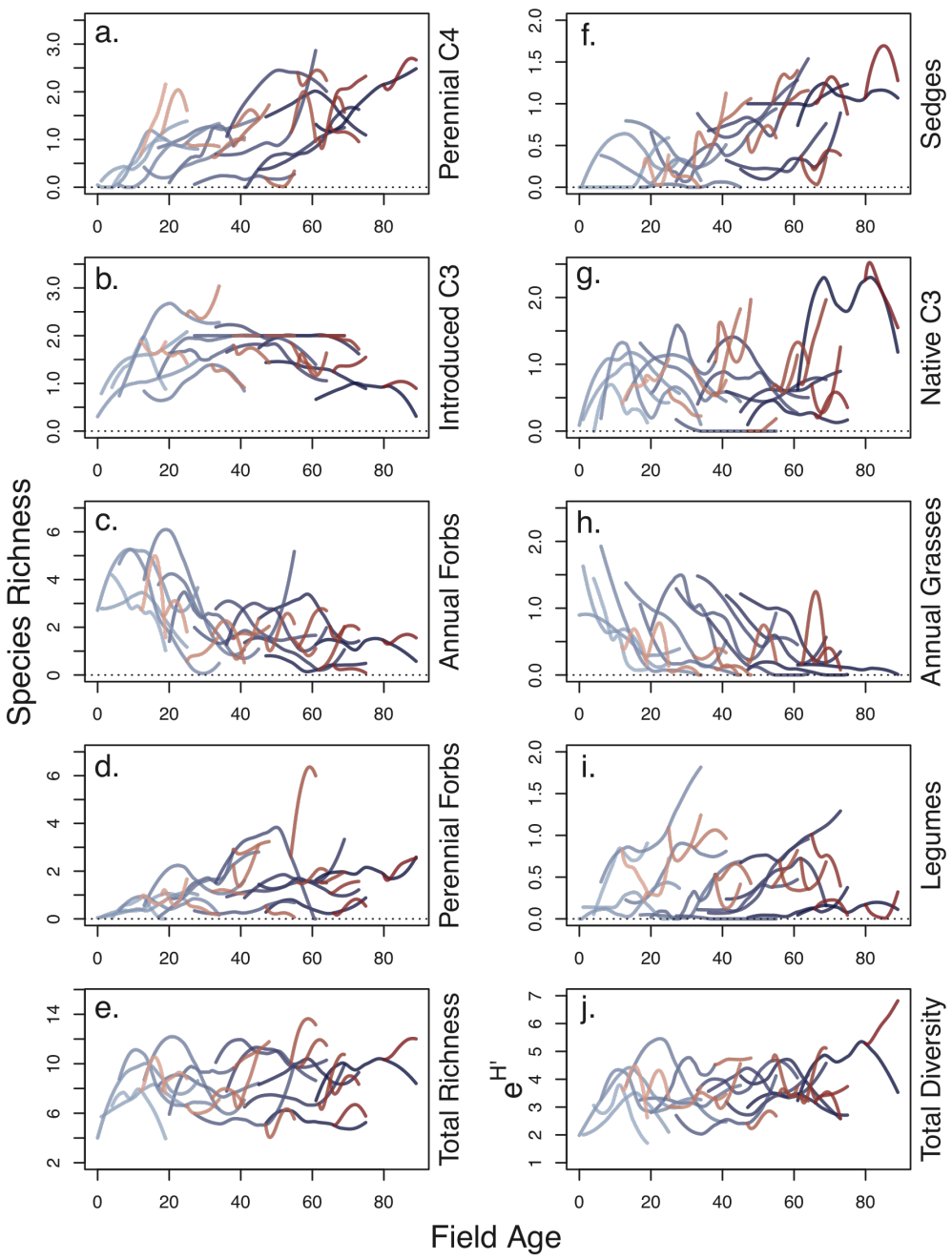

Figure S2: Changes in species richness **(a-i)** and Shannon diversity **(j)** as a function of field age (i.e. years since abandonment) for **(a-d,f-i)** various plant functional groups, and **(e,j)** all species. Each line shows the successional trajectory for a single field. Red and blue lines show dynamics in burned and unburned transects, respectively. Lighter colours show dynamics for younger fields. Note the difference in scale for the vertical axis among panels.

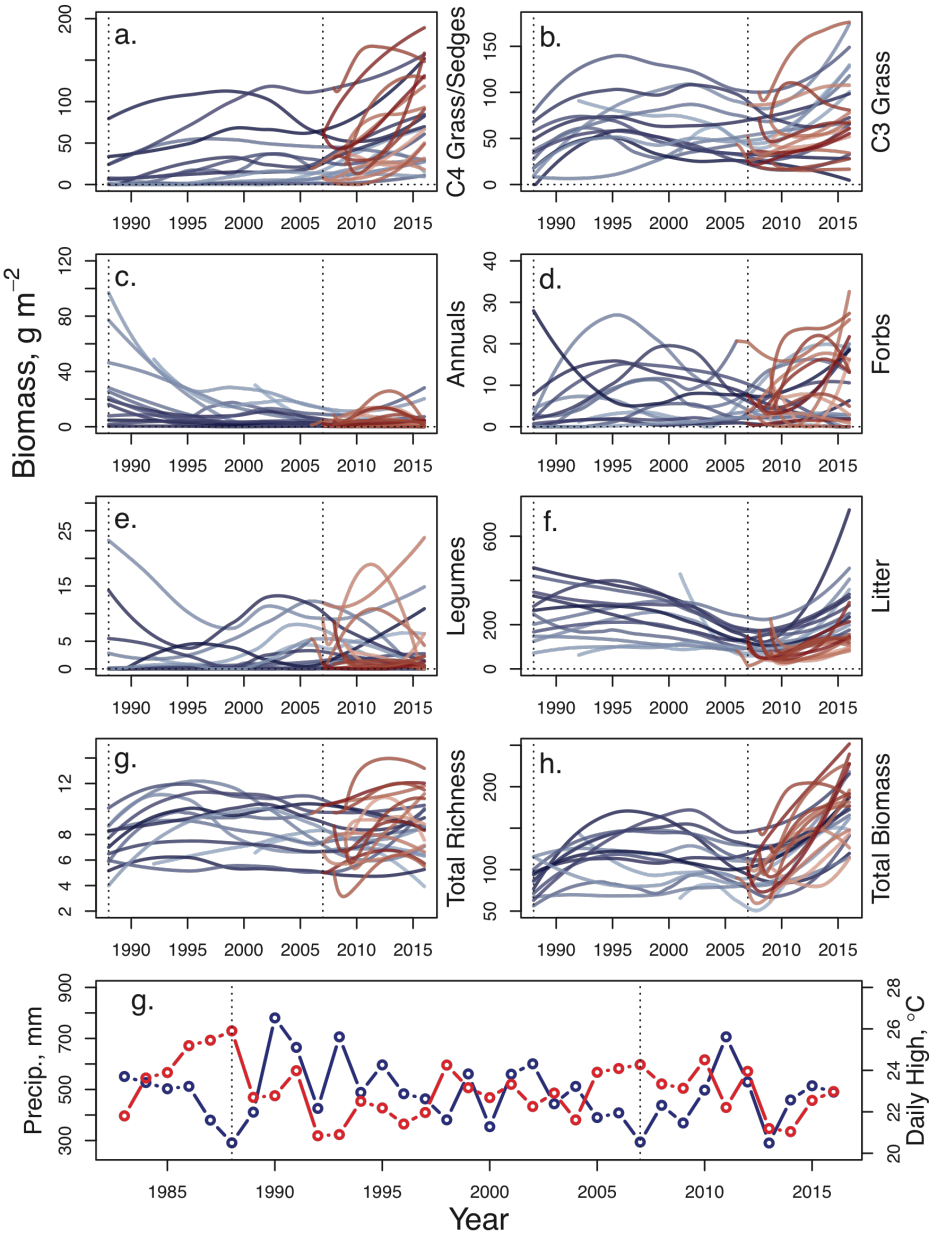

Figure S3: Changes in aboveground biomass **(a-f,h)** and species richness **(g)** as a function of time and weather for **(a-f)** various functional groups, and **(g-h)** all species. Colours and lines are as described in the legend for Fig. S2. **(g)** Time trends for mean daily maximum temperature (red) and total precipitation (blue), measured April-August. Dotted lines in all panels correspond to an extreme drought event in 1988 (the "North American Drought"), and a less severe drought in 2008. Note corresponding changes in species biomass and diversity.

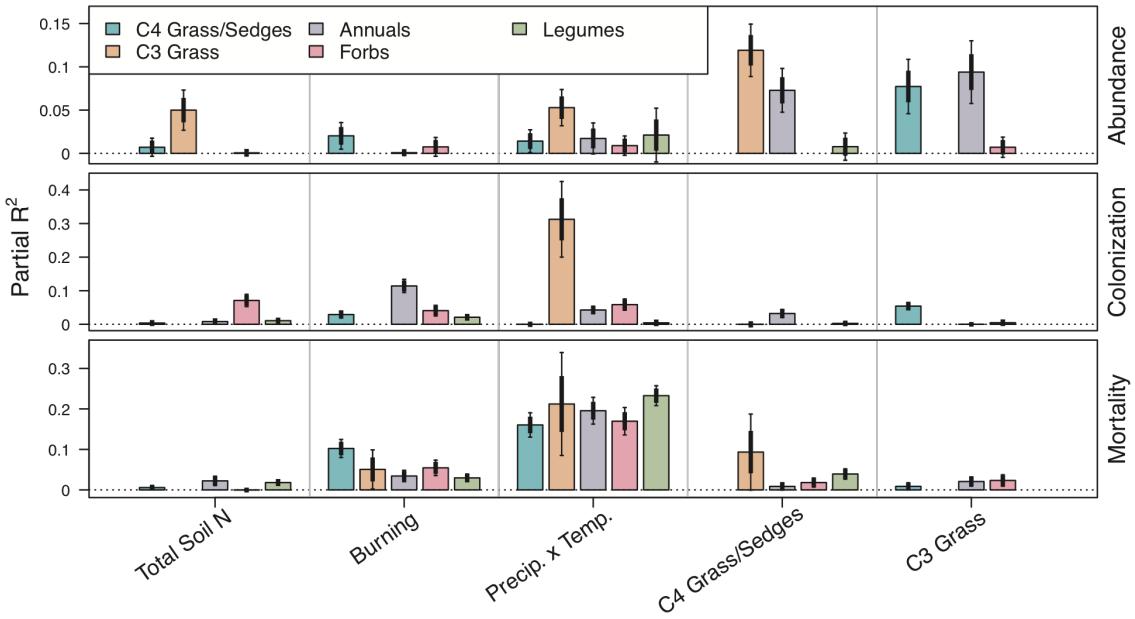

Figure S4: Partial R^2 showing relative explanatory power for parameters in the abundance, colonization, and mortality regressions. Bars and lines show mean ± one and two standard deviations, based on 1,000 nonparametric bootstrapped iterations (i.e. resampling data and predictions with replacement).

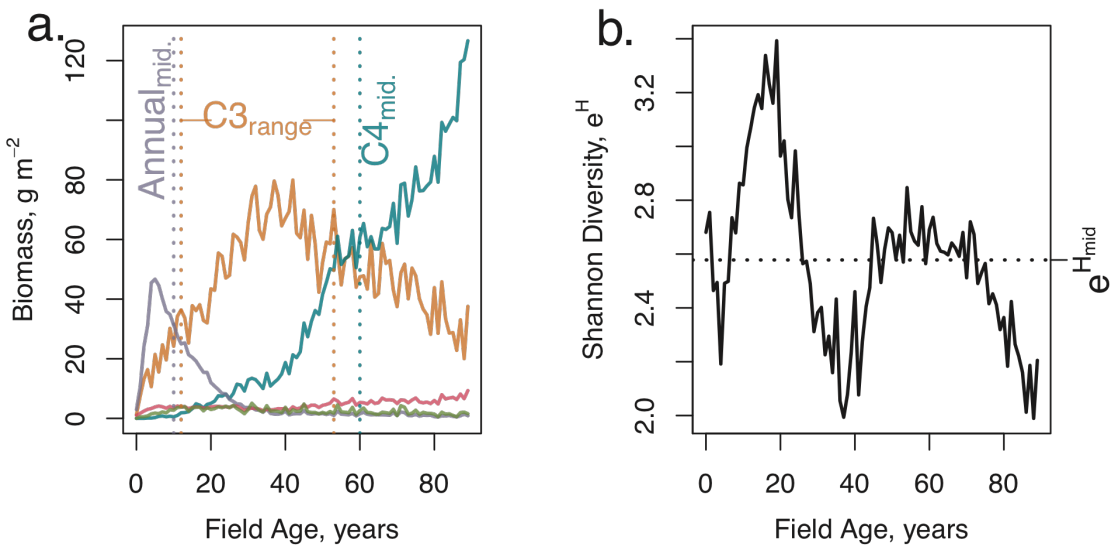

Fig. S5: Example output from one run of the metapopulation model (mean conditions used for all variables), demonstrating summary statistics. See text for details.

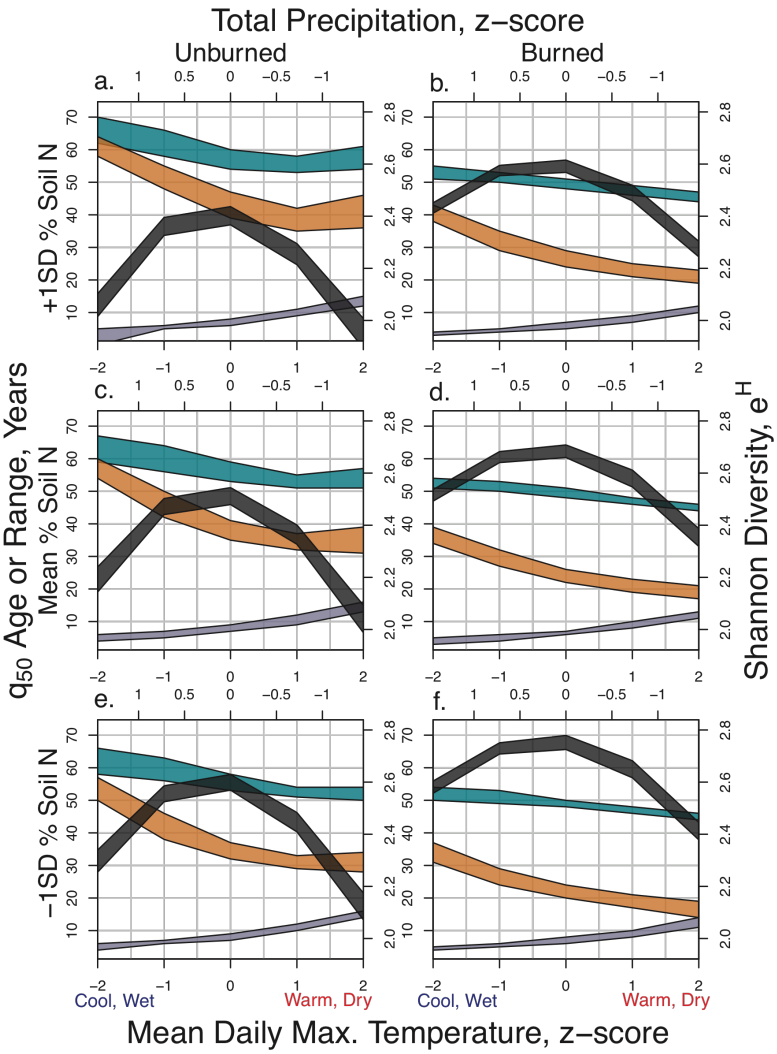

Figure S6: Results from the metapopulation models. Shaded regions show mean ± one standard deviation, calculated across 1,000 simulations. Purple shows $\text{Annual}_{\text{mid.}}$, light blue shows $\text{C4}_{\text{mid.}}$, orange shows C3_{range}, and black shows e^{H}_{mid}. Horizontal axes shows mean change in precipitation (top) and temperature (bottom) ranging from cool and wet years to warm and dry years, following the relationship described in Fig. 2b. Panels in the left column **(a,c,e)** and right column **(b,d,f)** show simulation results for unburned and burned plots, respectively. Panels from top row, middle row, and bottom row show scenarios with decreasing soil fertility **(a-b,c-d,e-f)**. See Fig. S5 for examples of how to interpret the indices.

Appendices for Chapter Two:
Harnessing Uncertainty to Approximate Mechanistic Models

A MacArthur's resource competition model

A.I Derivation of the model

MacArthurs 1970 resource competition model tracks dynamics in a community of n interacting species that jointly compete for a pool of l_{max} limiting resources (MacArthur, 1970). Resources in this model are perfectly substitutable, in that a sufficient quantity of any combination of resource types allows species to persist (Tilman, 1982). However, species differ in their ability to acquire each resource, and resources can differ in terms of how much they contribute to species growth rates (though the relative contribution of each resource is shared across all species). An example of such a resource might be energy, or food items of similar nutritious value (Chesson, 1990). Following the notation of Chesson (1990) (see note below), dynamics of the abundance of species i, X_i, follow:

$$\frac{1}{X_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} = b_i\left(\sum_{l=1}^{l_{max}} c_{il}w_l R_l - m_i\right) \tag{SA1}$$

where b_i is the growth rate of species i, c_{il} is the capture rate of resource l by species i, w_l describes the contribution of resource l to species growth relative to other resources, and m_i is the total amount of resource that must be acquired by a species per unit time in order to maintain zero population growth (*i.e.* a basal metabolic rate, or population mortality rate). For simplicity, we group the c_{il} terms into a matrix $\boldsymbol{c}$, where each species is represented by a row, and each resource by a column.

The resources in this model grow logistically in the absence of consumption, with dynamics of the availability of resource l, R_l, following:

$$\frac{1}{R_l}\frac{\mathrm{d}R_l}{\mathrm{d}t} = g_l(1-\frac{R_l}{\kappa_l}) - \sum_{j=1}^{n} c_{jl}X_j \tag{SA2}$$

where g_l is the growth rate of the resource, κ_l is its carrying capacity in the absence of consumers, and j is used to index species. MacArthur's great insight was to realize that if dynamics of the resource are much faster than dynamics of the consumer species, then we can assume that the resource is always at equilibrium relative to consumer abundance. Following this approximation:

$$R_l = \kappa_l(1-\sum_{j=1}^{n}\frac{c_{jl}}{g_l}X_j) \tag{SA3}$$

We can then substitute this approximation for R_l into Eq. (SA1), yielding a simplified model of the

form:

$$\frac{1}{X_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} = b_i\left(\sum_{l=1}^{l_{max}} c_{il}w_l\kappa_l - m_i - \sum_{j=1}^{n}\sum_{l=1}^{l_{max}} c_{il}c_{jl}\frac{w_l\kappa_l}{g_l}X_j\right) \tag{SA4}$$

In order to relate this to more typical formulations of Lotka-Volterra competition, we can group these parameters into two major terms:

$$a_{ij} = \sum_{l=1}^{l_{max}} c_{ij}c_{ji}\frac{w_l\kappa_l}{g_l} \tag{SA5}$$

$$k_i = \sum_{l=1}^{l_{max}} c_{il}w_l\kappa_l - m_i \tag{SA6}$$

where a_{ij} describes the effect of species j on the growth rate of species i, and k_i is total uptake of resource l by species i at equilibrium. The terms of Eq. (SA5) (which include the effect of species i on itself, a_{ii}) are typically grouped together into matrix $\boldsymbol{a}$, known as a "community" matrix (Dormann, 2008). Note that for this model, $\boldsymbol{a}$ is necessarily symmetric (*i.e.* $a_{ij} = a_{ji}$) (MacArthur, 1970).

Substituting the terms in Eqs. (SA5) and (SA6) into Eq. (SA4) gives:

$$\frac{1}{X_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} = b_i\left(k_i - \sum_{j=1}^{n} a_{ij}X_j\right) \tag{SA7}$$

which is equivalent to Eq. (2) in the main text. To make this system as simple as possible, we further reduce it from the form used in MacArthur (1970) and Chesson (1990) with a number of assumptions:

1. $w_l\kappa_l = 1$ 2. $g_l = 1$ 3. $m_i = 0.1$ 4. $b_i = 0.1$

Note that this meets MacArthurs assumption that $g_l \gg b_i$. In biological terms, these assumptions effectively require, respectively, that the the total nutritious value of each resource available in the absence of consumption, $w_l\kappa_l$, is constant across resources and species and is scaled to 1; growth rates of all resources, g_l, are equivalent across resources and are scaled to 1; species maintenance rates, m_i, are equivalent across species and are scaled to 0.1; and growth rates, b_i, are equivalent across species and scaled to equal 0.1 as well. Thus, we can further simplify Eqs. (SA5) and (SA6) as:

$$a_{ij} = \sum_{l=1}^{l_{max}} c_{il}c_{jl} \tag{SA8}$$

$$k_i = \sum_{l=1}^{l_{max}} c_{il} - 0.1 \tag{SA9}$$

Note that Eq. (SA7) differs from the form of the Lotka-Volterra system that is presented in many textbooks:

$$\frac{1}{X_i}\frac{\mathrm{d}X_i}{\mathrm{d}t} = r_i\left(1 - \frac{X_i + \sum_{j\neq i}\alpha_{ij}X_j}{K_i}\right) \tag{SA10}$$

While Eq. (SA9) is mathematically equivalent to Eq. (SA10), the parameterization differs somewhat. Specifically, α_{ij}, K_i, and r_i are usually meant to represent competition coefficients, carrying capacity, and relative growth rates, respectively, but all scaled by the inhibitory effect of species i on itself (*i.e.* a_{ii}). Thus, to convert the parameters we use here to these "classical" versions of the model requires $\alpha_{ij} = a_{ij}/a_{ii}$, $K_i = k_i/a_{ii}$, and $r_i = b_i k_i$ (Chesson, 2000). In MacArthur (1970), parameters a and k are written as α and K, respectively. To avoid potential confusion, we therefore use the notation and derivation from Chesson (1990) with three somewhat pedantic exceptions:

1. we use l_{max} rather than m to indicate the number of limiting resources
2. we use g_l rather than r_l to indicate the growth rate of the resource
3. we use κ_l rather than K_l to indicate the carrying capacity for the resource

We make these notational changes because m is also used to describe the minimum quantity of resource needed by each species, and in order to reserve r and K to indicate species growth rates and carrying capacities in the "textbook" Lotka-Volterra system.

A.II Model equilibria and stability

Following Chesson (1990), equilibrium population abundance in this model can be found by solving:

$$\vec{X}^* = \boldsymbol{a}^{-1}\vec{k} \tag{SA11}$$

where $\vec{X}^*$ is the vector of species equilibrium abundances, $\boldsymbol{a}^{-1}$ is the inverted form of matrix $\boldsymbol{a}$ (provided that it exists), and $\vec{k}$ is a vector containing the k_i elements from Eq. (SA8). If all values of $\vec{X}^*$ are positive, then the solution is said to be "feasible" (*i.e.* physically possible).

A necessary, but insufficient, criterion for an equilibrium to be stable in this system is that the number of coexisting species, n, must be less than or equal to the number of limiting resources, l_{max}, and the consumption vectors of each species, made up of the elements of $\boldsymbol{c}$, must be linearly independent from one another (MacArthur, 1970; Levin, 1970). Because the elements of $\boldsymbol{a}$ are simply the product of pairwise elements of $\boldsymbol{c}$, this implies that coexistence can only occur if $\boldsymbol{a}$ contains n linearly independent elements (*i.e.* $\boldsymbol{a}$ must be of "full rank") (Chesson, 1990). However, even if this condition is met, it is possible that the equilibrium is not locally stable.

To fully determine stability for the equilibria in this system, we therefore calculated the eigenvalues of the Jacobian matrix, composed of elements:

$$J_{ii} = b_i \left(k_i - 2a_{ii}X_i^* - \sum_{j \neq i} a_{ij}X_j^* \right) \tag{SA12}$$

$$J_{ij} = -b_i a_{ij} X_i^* \tag{SA13}$$

From Eq. (SA8), note that at equilibrium, provided that all species have positive abundance, $k_i = \sum_j a_{ij}X_j^*$, which implies that $J_{ii} = -b_i a_{ii} X_i^*$. However, if any species are competitively excluded, this simplification no longer holds, as $\sum_j a_{ij}X_j^*$ can be greater than k_i for these species.

If all eigenvalues of $\boldsymbol{J}$ are negative, this indicates that the solution is locally stable. Furthermore, though local equilibria are not always globally stable in general Lotka-Volterra systems (Goh, 1977), any feasible, locally stable equilibrium in the MacArthur model is also globally stable (Chesson, 1990; Haygood, 2002). Thus, if we are able to identify a unique equilibrium that was both feasible and locally stable, it will also be globally stable.

We offer one piece of caution in interpreting this assumption of global stability. The technical requirement for this form of stability is that the solution is "interior" – *i.e.* all resources and species have positive abundances at equilibrium (Haygood, 2002). Note, however, that in a system with $n > l_{max}$ coexisting species (or with fewer than n linearly independent consumption vectors), only subsets of species can coexist at equilibrium. It is possible that multiple such subsets will appear to be "globally stable," in the sense that they are stable in the absence of other species. For example, given two limiting resources and species A, B, and C, the full community of all three species may be unable to stably coexist, but all three pairwise communities (A and B; B and C; A and C) may be able to stably coexist in the absence of the third competitor. While this technically meets the criteria for "global stability of interior equilibria" from the perspective of each two species community living in isolation, calling such a system "globally stable" is potentially confusing, as there is no single mixture of species that will dominate under all circumstances.

Furthermore, because these species are likely to have consumption vector that are linearly related to one another, there is a strong possibility that the system is in fact "neutrally" stable with respect to mixtures of $n > l_{max}$ species. This can make identifying stable equilibria through the use of numerical techniques exceedingly difficult, and we therefore strongly suggest that the stability of equilibria in this system be determined analytically through the use of the Jacobian matrix.

B The covariance matrix

B.I Derivation of the covariance matrix

In the deterministic model:

$$a_{ij} = \sum_{l=1}^{l_{max}} c_{il}c_{jl} \tag{SB1}$$

We now add process noise to the capture rate of resource l by species i as:

$$\tilde{c}_{il} = c_{il} + \sigma Z_{il} \tag{SB2}$$

where $Z_{il} \sim \mathcal{N}(0,1)$, and $\tilde{c}_{ij}$ indicates a normally distributed random variable centered on c_{ij} with standard deviation determined by σ. We assume that $\{Z_{il}\}_{i=1,\ldots,n;l=1,\ldots,l_{max}}$ are independently distributed. This implies that $\mathrm{Cov}(Z_{il}, Z_{jl}) = 0$ and $\mathrm{Cov}(Z_{il}, Z_{im}) = 0$. Note (as discussed below) that the expected value of a function of $\tilde{c}_{ij}$ is not necessarily equal to the same function applied to c_{ij} (*i.e.* $\mathbb{E}(f(\tilde{c}_{ij})) \neq f(c_{ij})$).

Because of the symmetrical nature of $\boldsymbol{a}$ in the MacArthur model, the covariance matrix of relationships among all components of $\boldsymbol{a}$ can be fully described using just six classes of functions. For simplicity, we refer to these six types of terms using the following shorthand: $\mathrm{Var}(\tilde{a}_{ii})$, $\mathrm{Var}(\tilde{a}_{ij})$, $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji})$, $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii})$, $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ik})$, and $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{kl})$. More specifically:

1. $\mathrm{Var}(\tilde{a}_{ii})$ and 2. $\mathrm{Var}(\tilde{a}_{ij})$ describe, respectively, the variance in the effects of self-limitation (a_{ii}), and variance in the effects of interspecific competition (a_{ij})
3. $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji})$ describes the covariance between any pair of interaction terms where one term describes self-limitation of species i by species j, and the second describes self-limitation of species j by species i.
4. $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii})$ describes the covariance between any pair of interaction terms where one term describes self-limitation and the other describes either the effect of that self-limiting species on another species, or the effect of another species on that self-limiting species (*i.e.* this term also includes $\mathrm{Cov}(\tilde{a}_{ji}, \tilde{a}_{ii})$)
5. $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ik})$ shows the covariance between two interspecific competition terms that include three distinct species (*i.e.* this term also includes $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ki})$, $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{jk})$, etc).
6. $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{kl})$ describes covariance between any two terms where none of the species in one of the coefficients is included in the other coefficients (*i.e.* this term also includes $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{kk})$, $\mathrm{Cov}(\tilde{a}_{ii}, \tilde{a}_{jj})$, etc).

Analytical forms corresponding to these six classes of parameters and derivations for those forms (as well as expected values for $\tilde{\boldsymbol{a}}$), are listed below. Note that these analytical functions closely approximate the empirical estimates derived from simulations, even when n is large and $n \neq l$ (Fig. S1). The magnitudes of the six types of components in the covariance matrix also tend to cluster together, and, given some relatively broad conditions that we identify, have relative magnitudes $\mathrm{Var}(\tilde{a}_{ii}) > \mathrm{Var}(\tilde{a}_{ij}) = \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji}) > \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii}) > \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ik}) > \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{kl}) = 0$. (For the precise conditions needed to ensure that these inequalities hold, see Appendix B.II.)

For these derivations, it is helpful to recall the formula for the covariance of two random variables, X and Y, with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, respectively: $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$. Thus, the expected value of their product, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) + \mathrm{Cov}(X,Y)$. Similarly, recall that $\sigma_X^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Calculations of expected values

Note that because of the assumption that noise in each component of $\tilde{\boldsymbol{c}}$ is independent:

$$\mathbb{E}(\tilde{c}_{il}\tilde{c}_{jl}) = c_{il}c_{jl} \quad \text{for} \quad i \neq j \tag{SB.0.1}$$

Furthermore,

$$\mathbb{E}(\tilde{c}_{il}^2) = c_{il}^2 + \sigma^2 \tag{SB.0.2}$$

because $\mathrm{Var}(\tilde{c}_{ij}) = \mathrm{Var}(c_{ij} + \sigma Z_{il}) = \sigma^2$.

With $\tilde{a}_{ij} = \sum_{l=1}^{l_{max}} \tilde{c}_{il}\tilde{c}_{jl}$, we know that:

$$\begin{aligned}\mathbb{E}(\tilde{a}_{ij}) = \mathbb{E}(\sum_{l=1}^{l_{max}} \tilde{c}_{il}\tilde{c}_{jl}) = \sum_{l=1}^{l_{max}} \mathbb{E}(\tilde{c}_{il}\tilde{c}_{jl}) &= \sum_{l=1}^{l_{max}} \mathbb{E}(\tilde{c}_{il})\mathbb{E}(\tilde{c}_{jl}) \\ &= \sum_{l=1}^{l_{max}} c_{il}c_{jl} = a_{ij} \quad \text{for} \quad i \neq j\end{aligned} \tag{SB.0.3}$$

and

$$\mathbb{E}(\tilde{a}_{ii}) = \mathbb{E}(\sum_{l=1}^{l_{max}} \tilde{c}_{il}^2) = \sum_{l=1}^{l_{max}} \mathbb{E}\tilde{c}_{il}^2 = \sum_{l=1}^{l_{max}} (c_{il}^2 + \sigma^2) = a_{ii}^2 + l_{max}\sigma^2 \tag{SB.0.4}$$

Calculations of covariance classes

Class 1: $\boxed{\mathrm{Var}(\tilde{a}_{ii}) = 2\sigma^2(l_{max}\sigma^2 + 2a_{ii})}$ (SB.1)

Because of the independence assumption

$$\mathrm{Var}(\tilde{a}_{ii}) = \mathrm{Var}(\sum_{l=1}^{l_{max}} \tilde{c}_{il}^2) = \sum_{l=1}^{l_{max}} \mathrm{Var}(\tilde{c}_{il}^2) \tag{SB.1.1}$$

If $X = \mu + \sigma Z$, then

$$\begin{aligned}\mathrm{Var}X^2 &= \mathrm{Var}(\mu^2 + 2\mu\sigma Z + \sigma^2 Z^2) \\ &= \mathrm{Var}(2\mu\sigma Z + \sigma^2 Z^2) \\ &= \mathrm{Var}(2\mu\sigma Z) + \mathrm{Var}(\sigma^2 Z^2) + \mathrm{Cov}(2\mu\sigma Z, \sigma^2 Z^2) \\ &= 4\mu^2\sigma^2 + \sigma^4\mathrm{Var}(Z^2) + \mathbb{E}(2\mu\sigma^3 Z^3) - \mathbb{E}(2\mu\sigma Z)\mathbb{E}(\sigma^2 Z^2) \\ &= 4\mu^2\sigma^2 + 2\sigma^4\end{aligned} \tag{SB.1.2}$$

Hence,

$$\mathrm{Var}(\tilde{c}_{il}^2) = 2\sigma^4 + 4\sigma^2 c_{il}^2 \tag{SB.1.3}$$

and

$$\begin{aligned}\mathrm{Var}(\tilde{a}_{ii}) &= 2\sigma^2\left(\sum_{l=1}^{l_{max}}(\sigma^2+2c_{il}^2)\right) \\ &= 2\sigma^2\left(l_{max}\sigma^2+2\sum_{l=1}^{l_{max}}c_{il}^2\right) \\ &= 2\sigma^2(l_{max}\sigma^2+2a_{ii})\end{aligned} \tag{SB.1.4}$$

Lastly, note that if we assume $\sigma^4 \approx 0$ (*i.e.* $\sigma \ll 1$), then we can approximate Eq. (SB.1) as:

$$\mathrm{Var}(\tilde{a}_{ii}) \approx 4\sigma^2 a_{ii} \tag{SB.1.5}$$

This is the linearized approximation used in Fig. 1 in the main text.

Class 2: $$\boxed{\mathrm{Var}(\tilde{a}_{ij}) = \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj})} \tag{SB.2}$$

$$\mathrm{Var}(\tilde{a}_{ij}) = \mathrm{Var}\left(\sum_{l=1}^{l_{max}}(\tilde{c}_{il}\tilde{c}_{jl})\right) = \sum_{l=1}^{l_{max}}\mathrm{Var}(\tilde{c}_{il}\tilde{c}_{jl}) \tag{SB.2.1}$$

If X and Y are independent random variables,

$$\begin{aligned}\mathrm{Var}(X,Y) &= \mathbb{E}\left((XY)^2\right) - \mathbb{E}\,(XY)^2 \\ &= \mathbb{E}(X^2)\mathbb{E}(Y^2) - \mathbb{E}(X)^2\mathbb{E}(Y)^2 \\ &= \mathbb{E}(X^2)\mathbb{E}(Y^2) - \mathbb{E}(X^2)\mathbb{E}(Y)^2 + \mathbb{E}(X^2)\mathbb{E}(Y)^2 - \mathbb{E}(X)^2\mathbb{E}(Y)^2 \\ &= \mathbb{E}(X^2)\mathrm{Var}(Y) + \mathbb{E}(X^2)\mathbb{E}(Y)^2 - \mathbb{E}(X)^2\mathbb{E}(Y)^2 \\ &= \mathbb{E}(X^2)\mathrm{Var}(Y) + \mathbb{E}(Y)^2\mathrm{Var}(X) + \mathbb{E}(Y^2)\mathrm{Var}(X) - \mathbb{E}(Y^2)\mathrm{Var}(X) \\ &= \mathbb{E}(X^2)\mathrm{Var}(Y) + \mathbb{E}(Y^2)\mathrm{Var}(X) - \mathrm{Var}(X)\mathrm{Var}(Y)\end{aligned} \tag{SB.2.2}$$

Hence,

$$\begin{aligned}\mathrm{Var}(\tilde{c}_{il}\tilde{c}_{jl}) &= \mathbb{E}(\tilde{c}_{il}^2)\mathrm{Var}(\tilde{c}_{jl}) + \mathbb{E}(\tilde{c}_{jl}^2)\mathrm{Var}(\tilde{c}_{il}) - \mathrm{Var}(\tilde{c}_{il})\mathrm{Var}(\tilde{c}_{jl}) \\ &= \sigma^2\mathbb{E}(\tilde{c}_{il}^2) + \sigma^2\mathbb{E}(\tilde{c}_{jl}^2) - \sigma^4\end{aligned} \tag{SB.2.3}$$

If $X = \mu + \sigma Z$, then

$$\mathbb{E}(X^2) = E(\mu^2 + 2\mu\sigma Z + \sigma^2 Z^2) = \mu^2 + \sigma^2 \tag{SB.2.4}$$

Hence,

$$\begin{aligned}\sigma^2\mathbb{E}(\tilde{c}_{il}^2) + \sigma^2\mathbb{E}(\tilde{c}_{jl}^2) - \sigma^4 &= \sigma^2(c_{il}^2+\sigma^2) + \sigma^2(c_{jl}^2+\sigma^2) - \sigma^4 \\ &= \sigma^2(c_{il}^2 + c_{jl}^2) + \sigma^4\end{aligned} \tag{SB.2.5}$$

Therefore,

$$\begin{aligned}\mathrm{Var}(\tilde{a}_{ij}) &= \sum_{l=1}^{l_{max}} \mathrm{Var}(\tilde{c}_{il}\tilde{c}_{jl}) \\ &= \sum_{l=1}^{l_{max}} \left(\sigma^2(\tilde{c}_{il}^2 + \tilde{c}_{jl}^2) + \sigma^4\right) \\ &= \sigma^2\left(l_{max}\sigma^2 + \sum_{l=1}^{l_{max}}(c_{il}^2 + c_{jl}^2)\right) \\ &= \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj})\end{aligned} \tag{SB.2.6}$$

Lastly, note that if we assume $\sigma^4 \approx 0$ (*i.e.* $\sigma \ll 1$), then we can approximate Eq. (SB.3) as:

$$\mathrm{Var}(\tilde{a}_{ij}) \approx \sigma^2(a_{ii} + a_{jj}) \tag{SB.2.7}$$

This is the linearized approximation used in Fig. 1 in the main text.

Class 3: $$\boxed{\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji}) = \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj})} \tag{SB.3}$$

Because $\boldsymbol{a}$ is symmetrical, we know that a_{ij} is equal to a_{ji}, as are any individual observations of $\tilde{a}_{ij}$ and $\tilde{a}_{ji}$. Thus,

$$\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji}) = \mathrm{Var}(\tilde{a}_{ij}) = \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj}) \tag{SB.3.1}$$

Class 4: $$\boxed{\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii}) = 2\sigma^2 a_{ij}} \tag{SB.4}$$

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii}) &= \mathrm{Cov}\left(\sum_{l=1}^{l_{max}}(\tilde{c}_{il}, \tilde{c}_{jl}), \sum_{k=1}^{l_{max}} \tilde{c}_{ik}^2\right) \\ &= \sum_{l=1}^{l_{max}}\sum_{k=1}^{l_{max}} \mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl}, \tilde{c}_{ik}^2) \\ &= \sum_{l=1}^{l_{max}} \mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl}, \tilde{c}_{il}^2)\end{aligned} \tag{SB.4.1}$$

We need to calculate $\mathrm{Cov}(XY, X^2)$, where X and Y are independent random variables.

$$\begin{aligned}\mathrm{Cov}(XY, X^2) &= \mathbb{E}(X^3, Y) - \mathbb{E}(XY)\mathbb{E}(X^2) \\ &= \mathbb{E}(X^3)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y)\mathbb{E}(X^2) \\ &= \left(\mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2)\right)\mathbb{E}(Y)\end{aligned} \tag{SB.4.1}$$

Defining X as $X = \mu + \sigma Z$, and simplifying the first term on the right hand side of the equation,

$$\begin{aligned}\mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2) &= \mathbb{E}\left(\mu^3 + 3\mu^2\sigma Z + 3\mu\sigma^2 Z^2 + \sigma^3 Z^3 - \mu(\mu^2 + \sigma^2)\right) \\ &= \mu^3 + 3\mu\sigma^2 - \mu^3 - \mu\sigma^2 = 2\mu\sigma^2\end{aligned} \tag{SB.4.2}$$

Hence,

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ii}) &= \sum_{l=1}^{l_{max}} \mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl},\tilde{c}_{il}^2) \\ &= \sum_{l=1}^{l_{max}} (2\sigma^2 c_{il}c_{jl}) = 2\sigma^2 a_{ij}\end{aligned} \tag{SB.4.3}$$

Class 5:

$$\boxed{\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) = \sigma^2 a_{jk}} \tag{SB.5}$$

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) &= \mathrm{Cov}(\sum_{l=1}^{l_{max}} (\tilde{c}_{il}\tilde{c}_{jl}), \sum_{m=1}^{l_{max}} \tilde{c}_{im}\tilde{c}_{km}) \\ &= \sum_{l=1}^{l_{max}}\sum_{m=1}^{l_{max}} \mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl},\tilde{c}_{im}\tilde{c}_{km}) \\ &= \sum_{l=1}^{l_{max}} \mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl},\tilde{c}_{il}\tilde{c}_{kl})\end{aligned} \tag{SB.5.1}$$

If X, U, and V are independent random variables, then

$$\begin{aligned}\mathrm{Cov}(XU,XV) &= \mathbb{E}(X^2UV) - \mathbb{E}(XU)\mathbb{E}(XV) \\ &= \mathbb{E}(X^2)\mathbb{E}(U)\mathbb{E}(V) - \mathbb{E}(X)^2\mathbb{E}(U)\mathbb{E}(V) \\ &= \mathrm{Var}(X)\mathbb{E}(U)\mathbb{E}(V)\end{aligned} \tag{SB.5.2}$$

Hence,

$$\begin{aligned}\mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl},\tilde{c}_{il}\tilde{c}_{kl}) &= \mathrm{Var}(\tilde{c}_{il})\mathbb{E}(\tilde{c}_{jl})\mathbb{E}(\tilde{c}_{kl}) \\ &= \sigma^2 c_{jl}c_{kl}\end{aligned} \tag{SB.5.3}$$

Therefore,

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) &= \sum_{l=1}^{l_{max}} \mathrm{Cov}(\tilde{c}_{il}\tilde{c}_{jl},\tilde{c}_{il}\tilde{c}_{kl}) \\ &= \sum_{l=0}^{l_{max}} (\sigma^2 c_{jl}c_{kl}) = \sigma^2 a_{ik}\end{aligned} \tag{SB.5.4}$$

Class 6:

$$\boxed{\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}kl) = 0} \tag{SB.6}$$

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{kl}) &= \mathrm{Cov}\left(\sum_{h=1}^{l_{max}} \tilde{c}_{ih}\tilde{c}_{jh}, \sum_{g=1}^{l_{max}} \tilde{c}_{kg}\tilde{c}_{lg}\right) \\ &= \sum_{h=1}^{l_{max}}\sum_{g=1}^{l_{max}} \mathrm{Cov}\left(\tilde{c}_{ih}\tilde{c}_{jh},\tilde{c}_{kg}\tilde{c}_{lg}\right) = 0\end{aligned} \tag{SB.6.1}$$

because of the independence assumption.

B.II Relative values of the covariance matrix

Given that some relatively general criteria are met, the relative magnitudes of the components of the covariance matrix of $\tilde{\boldsymbol{a}}$ are related in the following manner (Fig. S1):

Claim: $\mathrm{Var}(\tilde{a}_{ii}) \overset{\text{(i)}}{>} \mathrm{Var}(\tilde{a}_{ij}) \overset{\text{(ii)}}{=} \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji})$
$\overset{\text{(iii)}}{>} \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii}) \overset{\text{(iv)}}{>} \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ik})$
$\overset{\text{(iv)}}{>} \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{kl}) \overset{\text{(vi)}}{=} 0$

Here, we derive the six specific criteria required to assure this specific of inequalities.

(i) $\mathrm{Var}(\tilde{a}_{ii}) > \mathrm{Var}(\tilde{a}_{ij})$ if $l_{max}\sigma^2 > a_{jj} - 3a_{ii}$ for all $i \neq j$

$$\mathrm{Var}(\tilde{a}_{ii}) = 2\sigma^2(l_{max}\sigma^2 + 2a_{ii})$$
$$\mathrm{Var}(\tilde{a}_{ij}) = \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj})$$

Hence,

$$\begin{aligned}\mathrm{Var}(\tilde{a}_{ii}) - \mathrm{Var}(\tilde{a}_{ij}) &= 2\sigma^2(l_{max}\sigma^2 + 2a_{ii}) - \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj}) \\ &= \sigma^2(2l_{max}\sigma^2 + 4a_{ii} - l_{max}\sigma^2 - a_{ii} - a_{jj}) \\ &= \sigma^2\left(l_{max}\sigma^2 + 4a_{ii} - (a_{ii} + a_{jj})\right) > 0 \\ \text{if} \quad l_{max}\sigma^2 &> a_{ii} + a_{jj} - 4a_{ii} = a_{jj} - 3a_{ii}\end{aligned}$$

(ii) $\mathrm{Var}(\tilde{a}_{ij}) = \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji})$

This is Class 3 from Appendix B.I

(iii) $\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji}) > \mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii})$ for all $i \neq j$

$$\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ji}) = \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj})$$
$$\mathrm{Cov}(\tilde{a}_{ij}, \tilde{a}_{ii}) = 2\sigma^2 a_{ij}$$

Hence,

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ji}) - \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ii}) &= \sigma^2(l_{max}\sigma^2 + a_{ii} + a_{jj} - 2a_{ij}) \\ &= \sigma^2(l_{max}\sigma^2 + \sum_{l=1}^{l_{max}} c_{il}^2 + \sum_{l=1}^{l_{max}} c_{jl}^2 - 2\sum_{l=1}^{l_{max}} c_{il}c_{jl}) \\ &= \sigma^2(l_{max}\sigma^2 + \sum_{l=1}^{l_{max}} \left(c_{il}^2 - c_{il}c_{jl} + c_{jl}^2\right)) \\ &= \sigma^2(l_{max}\sigma^2 + \sum_{l=1}^{l_{max}} \left(c_{il} + c_{jl}\right)^2) > 0\end{aligned}$$

(iv) $\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ii}) > \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik})$ if $2a_{ij} > a_{kj}$

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ii}) &= 2\sigma^2 a_{ij} \\ \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) &= \sigma^2 a_{jk}\end{aligned}$$

Hence,

$$\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ii}) - \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) = \sigma^2(2a_{ij} - a_{jk}) > 0 \quad \text{if } 2a_{ij} > a_{jk}$$

(v) $\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) > \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{kl})$ if $c_{jl}c_{kl} > 0$ for at least one l

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) &= \sigma^2 a_{jk} \\ \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{kl}) &= 0\end{aligned}$$

Hence,

$$\begin{aligned}\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{ik}) - \mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{kl}) &= \sigma^2 a_{jk} \\ &= \sigma^2 \sum_{l=1}^{l_{max}} (c_{jl}c_{kl}) > 0 \\ \text{if} \quad c_{jl}c_{kl} &> 0 \text{ for some } l\end{aligned}$$

(vi) $\mathrm{Cov}(\tilde{a}_{ij},\tilde{a}_{kl}) = 0$

This is Class 6 from Appendix B.I
The sequence of inequalities therefore holds if:

(i) $l_{max}\sigma^2 > a_{jj} - 3a_{ii}$, *i.e.* $\frac{1}{3} - \frac{l_{max}\sigma^2}{3a_{ii}} < \frac{a_{jj}}{a_{ii}} < 3 + \frac{l_{max}\sigma^2}{a_{ii}}$

(iv) $2a_{ij} > a_{kj}$, *i.e.* $\frac{1}{2} < \frac{a_{ij}}{a_{kj}} < 2$

(v) $a_{jk} > 0$, *i.e.* $c_{jl}c_{kl} > 0$ for some l

Biologically, these requirements can be interpreted as meaning that: **(i)** differences in self-

limitation among species are reasonably small; **(iv)** per-capita effects of species j on species i and k are of a similar magnitude; and **(v)** the species are both limited by at least one shared resource.

C Model simulation

We simulated three forms of the MacArthur model with process noise. First, we simulated the mechanistic model itself, using the paramteres sampled from $\tilde{\boldsymbol{c}}$ with added noise to directly calculate $\boldsymbol{a}$. Second, we simulated an "uncorrelated" model, in which we accounted for the variance components in $\tilde{\boldsymbol{a}}$ that arose from $\tilde{\boldsymbol{c}}$, but not for covariance among the elements of $\tilde{\boldsymbol{a}}$. Lastly, we simulated a "covariance" model, which accounted for the full variance-covariance relationship among elements of $\tilde{\boldsymbol{a}}$, as calculated in Appendix B.I.

For the tests described in the main text, we completed 20,000 iterations of each model, along a gradient of community sizes ($n = 2, 4, 6, 8, 10$) and process noise magnitudes ($\sigma = 0.01, 0.05, 0.1, 0.15$). For each iteration, we generated a unique set of consumption parameters for each species and resource following a Sobol sequence with Owen and Faure-Tezuka scrambling, using the `randtoolbox` package in the `R` programming language (Dutang and Savicky, 2015; R Core Team, 2016). This procedure produced a set of over-dispersed pseudo-random values, which increased the dissimilarity among species foraging vectors, and therefore increased the probability of coexistence (Chesson, 1990; Haygood, 2002). To further increase the likelihood of coexistence, we imposed a tradeoff by scaling consumption vectors for each species to have a mean value of one, which ensured that large values of c_{il} for one resource (*i.e.* effective foraging) were compensated by small values (*i.e.* ineffective foraging). As described in Appendix A.I, to reduce model complexity we set $g_l = m_i = 0.1$ for all species, and set $\kappa_l = g_l = 1$ for all resources.

After generating $\boldsymbol{c}$, we added process noise using a Gaussian distribution with variance σ^2 centered on the original $\boldsymbol{c}$ values, and truncated at zero. In general, σ was sufficiently small relative to the elements of $\boldsymbol{c}$ that this truncation did not greatly effect model outcomes (see Fig. 1a in the main text). To generate the mechanistic model estimate of $\boldsymbol{a}$, we simply followed the relationship in Eq. (SA5). To calculate the uncorrelated model estimate of $\boldsymbol{a}$, we sampled values from a Gaussian distribution centered on the expected values described in Eqs. (SB.0.1)-(SB.0.2), with variance determined by the equations in Class 1 and 2 of Appendix B.I (again, truncated at zero). Lastly, to calculate the covariance model estimate of $\boldsymbol{a}$, we sampled from a multivariate Gaussian distribution (again, truncated at zero) parameterized using the expected values and full covariance relationship described in Appendix B.I, using the `mvtnorm` package in `R` (Mi et al., 2009).

For each of these estimates of $\boldsymbol{a}$, we used the equations in Appendix A.II to calculate all potential equilibria, and to test the stability of those equilibria. For example, for mixtures that included species A, B, and C, we tested for feasibility and stability of the three species community, as well

as of all single-species and two-species communities. If multiple communities that were subsets of one another were stable, we included information only from the most diverse mixture (*e.g.* if a community of species A and B was stable in the absence of C, and a community that included all three species was stable, then we used only the estimate that included all three species). We then used the estimates of stable species abundances to calculate species richness, the Shannon diversity index, and the Simpson diversity index for each type of model (though we present results only for richness, because the other indices produced qualitatively similar results). As described in the main text, to simplify tracking the provenance of error in these models, we assumed perfect knowledge of k_i and b_i for each species.

Lastly, we compared the ability of the uncorrelated and the covariance models to accurately predict results from the mechanistic model. If multiple stable communities were predicted by a model (which was relatively common for cases of $n > l_{max}$), we compared mean estimates of the diversity statistics from across all of these stable communities (though after excluding communities that were subsets of one another, as outlined above). To test predictions of abundance, we compared estimates for all stable communities predicted by each model. In cases where one model predicted that a community was stable when another did not, we set abundance estimates to zero for all species in the unstable case.

D Methods for empirical data

D.I Removing eigenvalues from interaction matrix

Random noise will often increase a matrix that is not of full rank to one that is of full rank (*i.e.* it make all of the rows and columns of the matrix linearly independent, even if they were not before the noise was added) (Bates and Maechler, 2010). This means that the rank predicted by the uncorrelated model frequently differs from that in the mechanistic model. This is because, no matter how much noise is added to $\boldsymbol{c}$, mechanistic derivations of $\boldsymbol{a}$ can never have more than l_{max} linearly independent rows and columns (Chesson, 1990).

However, if the elements of $\boldsymbol{a}$ are calculated independently (*e.g.* through field measurements, or using the uncorrelated model methods described above), then even infinitesimal noise will frequently lead to linear independence among all rows and columns of $\tilde{\boldsymbol{a}}$ (Anderson et al., 2010; Bates and Maechler, 2010), and could thereby lead to predictions of coexistence among all competing species, even if $n \gg l_{max}$ (*n.b.* matrix $\tilde{\boldsymbol{a}}$ is technically a Gaussian Wigner matrix).

If the noise is small, it may be possible to identify these "extra" dimensions that were added by noise based on the magnitude of their eigenvalues, which will usually be smaller than that of the "true" eigenvalues (see Fig. 4 in the main text). This is similar to what happens to a sheet of paper

that is slightly crinkled - the paper will not be perfectly flat, but it will be *almost* flat. Given an n by n square matrix $\boldsymbol{M}$ with eigenvalues $\vec{e} = [e_1, e_2, \ ..., e_n]$ and eigenvectors $\boldsymbol{v} = [\vec{v}_1, \vec{v}_2, \ ..., \vec{v}_n]$, we can transform $\boldsymbol{M}$ such that $e_i = 0$ through the function:

$$\boldsymbol{M}_{new} = \boldsymbol{M} - e_i(\vec{v}_i\vec{v}_i^T)/(||\vec{v}_i||^2)$$

where $\vec{v}_i^T$ is the transpose of $\vec{v}_i$, and $||\vec{v}_i||^2$ is the square of the norm of $\vec{v}_i$ (*i.e.* the sum of the squared elements of $\vec{v}_i$) (Leskovec et al., 2014).

We can interpret this transformation as having "removed" the extra dimensions associated with that eigenvalue, and $\boldsymbol{M}_{new}$ will be of one unit lower rank than $\boldsymbol{M}$ (*i.e.* it will have one fewer set of linearly independent rows and columns). However, this transformation will not completely remove the effects of noise: if $\boldsymbol{M}$ was created by adding noise to a lower-dimensional matrix $\boldsymbol{N}$, matrix $\boldsymbol{M}_{new}$ may be more similar to $\boldsymbol{N}$ than $\boldsymbol{M}$ is, but $\boldsymbol{N}$ and $\boldsymbol{M}_{new}$ will not be identical. Furthermore, this process can add imaginary components to $\boldsymbol{M}_{new}$ even if they are not present in $\boldsymbol{M}$, and rounding errors can result in the re-addition of some removed eigenvalues if the procedure is repeated multiple times to remove more than one eigenvalue.

D.II Fitting the covariance model

To demonstrate that the covariance model can be fit to observed data resulting from the mechanistic MacArthur model, we simulated 100 observations of species abundances from this model under two scenarios. The first included three species and two limiting resources, and the second included three species and three limiting resources. In all cases, we used a fixed level of process noise, with $\sigma = 0.1$. For each of 5000 iterations, we then generated new estimates of $\boldsymbol{a}$ by sampling from the analytical covariance relationship, centered around the true expected values of $\boldsymbol{a}$. We then calculated species abundance distributions, the likelihood of the 100 observations given the covariance model, and the eigenvalues of the estimated $\boldsymbol{a}$ for each of these 5000 iterations. To estimate model likelihood, we simulated 10,000 estimated species abundance distributions for each potential value of $\boldsymbol{a}$, and used a nonparametric density smoother to calculate the probability density of the observed species abundances given the simulated distribution.

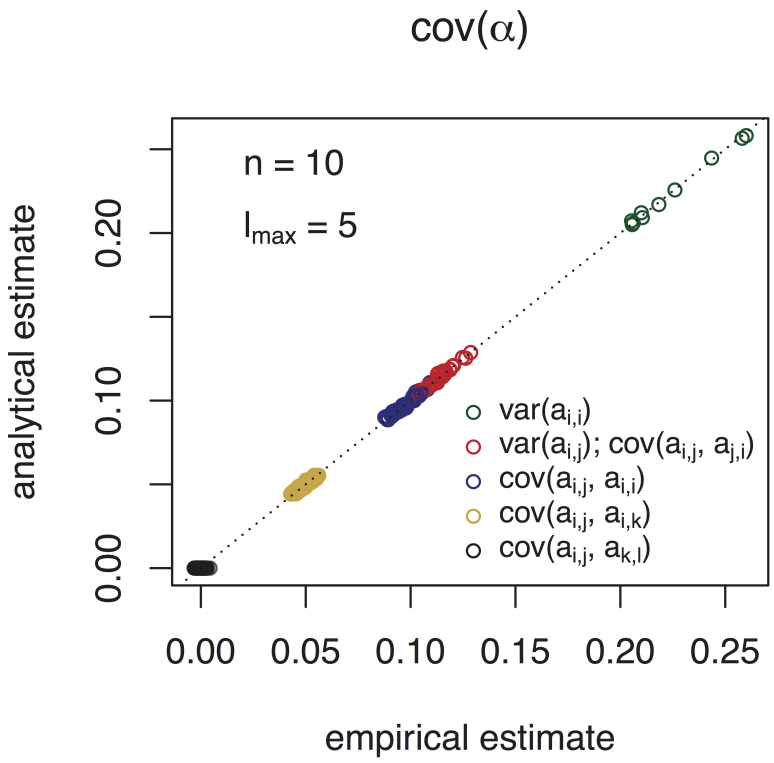

Figure S1: Comparison of analytical estimates of the covariance among terms in the interaction matrix $\tilde{\boldsymbol{a}}$, and empirical estimates generated from the mean of 20,000 simulations of a system with $n = 10$, $l_{max} = 5$, and $\sigma = 0.1$. Colors distinguish among different classes of variance and covariance, which cluster roughly into five groups, corresponding to the six functions discussed in Appendix B.I. See text for details.

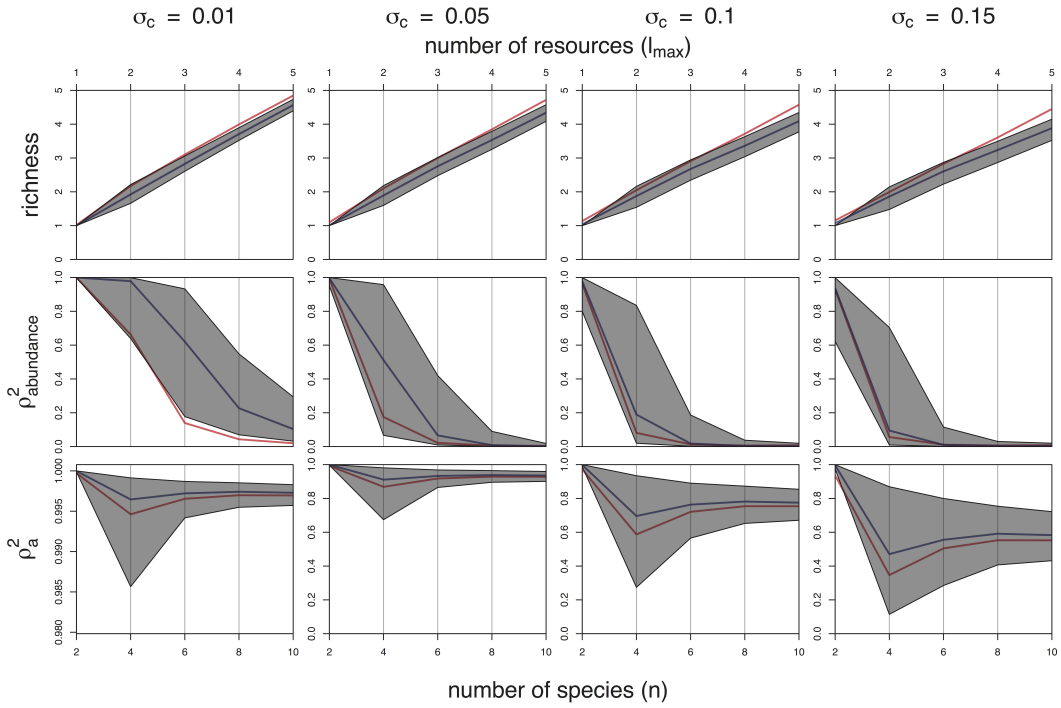

Figure S2: Correspondence between mechanistic model and Lotka-Volterra approximations as a function of process noise, σ_c, for communities of 2 to 10 species (note that $l_{max} = n/2$ in all cases). This figure is identical to Fig. 3 in the main text, except that it shows among-simulation variability in the mechanistic model rather than in the uncorrelated and covariance models. Solid lines show mean results for the uncorrelated model (red) and the covariance model (blue). Black shaded region shows variability in predictions from the underlying mechanistic model, as mean ± one standard deviation based on 20,000 simulations.

Appendices for Chapter Three:

Diversity, high dimensional tradeoffs, and trait evolution buffer effects of reduced resource limitation on coexistence

Appendix A: Resource competition model

Appendix A.I: Model derivation

We begin with a simple, linearized resource competition model, tracking the dynamics of n species competing for k limiting resources. Dynamics of each species are determined by:

$$dN_i/dt\ (1/N_i) = \min[r_{i,j}(R_j - R_{i,j}^*)] \quad \text{Eq. (SA.1)}$$

where N_i is the abundance of species i, $r_{i,j}$ describes the effect of resource j on the per-unit-biomass growth rate of species i, R_j is the availability of resource j, and $R_{i,j}^*$ is the minimum concentration of resource j at which species i experiences positive growth. Resource dynamics are modeled as responding instantaneously to species abundance, as:

$$R_j = R_{max,j} - \Sigma_i(q_{i,j}N_i) \quad \text{Eq. (SA.2)}$$

where $R_{max,j}$ is the maximum local concentration of resource j in the absence of consumption, and $q_{i,j}$ is the tissue resource concentration of resource j in species i.

Substituting Eq. (SA.1b) into (SA.1b) yields:

$$dN_i/dt\ (1/N_i) = \min[r_{i,j}(R_{max,j} - \Sigma_i(q_{i,j}N_i) - R_{i,j}^*)] \quad \text{Eq. (SA.3)}$$

We can solve this system at single-species equilibrium as:

$$N_i^* = \max[0, \min[(R_{max,j} - R_{i,j}^*)/q_{i,j}]] \quad \text{Eq. (SA.4)}$$

Given a single limiting resource, the species with the lowest R^* for that resource will competitively exclude all other species (Levin 1970; Tilman 1982). However, in systems with multiple limiting resources and multiple competing species, equilibrium outcomes depend on the precise ratios of R^* and R_{max} values among species and resources, as discussed in the main text (Tilman 1982).

To reduce the number of parameters in this system, we make two simplifying assumptions. First, we set $r_{i,j} = 1$ for all species and resources. Second, we assume optimal foraging *sensu* Tilman (1982) such that species tissue resource concentrations

(and thus resource uptake per unit biomass) are proportional to their R^* values for each resource (i.e. $q_{i,j} = w_i R_{i,j}^*$). This ensures that species consume resources in the same proportion by which they are limited by each resource, where w_i is a constant transforming units of R^* (resource mass / unit area) into units of tissue nitrogen concentration (resource mass / unit biomass). Thus, w_i is in units of (unit area / unit biomass), and describes the amount of area needed to produce a fixed amount of biomass in monoculture (given constant R_{max}). Thus, if we change units of area by a factor m (e.g. redefining the system in terms of square centimeters rather than square meters), w_i changes by a factor of $1/m$. For simplicity, we choose $w_i = 1$ for all species in our simulations.

Appendix A.II: Equilibrium conditions in monoculture

As noted in the main text, we solve for equilibria in multi-species systems either through comparison of species isoclines (for systems with two limiting resources), or through computational integration (for systems with three limiting resources). However, as noted in Eq. (SA.4), equilibrium biomass in monoculture can be calculated analytically. Including the simplifying assumptions described in *Appendix A.I*, we can write this as:

$$N_i^* = \max[0, \min[(R_{max,j} - R_{i,j}^*)/R_{i,j}^*]] \quad \text{Eq. (SA.7a)}$$

$$N_i^* = \max[0, \min[(R_{max,j}/R_{i,j}^* - 1)]] \quad \text{Eq. (SA.7b)}$$

Given a linear tradeoff among R^* values and a supply gradient that includes all potential ratios of resources with equal frequency (e.g. the tradeoffs and supply gradients that we use in the main text), we can show that the integral of N_i^* taken across all supply points is equal across all species. As an example, consider a system with two limiting resources, X, and Y. In this case:

$$N_i^* = \max[0, \min[(R_{max,X}/R_{i,X}^* - 1), ((m - R_{max,X})/(k - R_{i,X}^*) - 1)]] \quad \text{Eq. (SA.8)}$$

where k is a constant defining the linear tradeoff, such that $\Sigma_j R_{i,j}^* = k$ for all species i and resources j, and m is a constant defining the supply gradient, such that $\Sigma_j R_{max,j} = m$.

To identify when each resource is limiting, we can solve for the supply point $R_{max,X}$ that leads to equal limitation by resource X and Y:

$$(R_{max,X}/R_{i,X}^* - 1) = ((m - R_{max,X})/(k - R_{i,X}^*) - 1) \quad \text{Eq. (SA.9a)}$$

$(R_{max,X}/R_{i,X}^*) = (m - R_{max,X})/(k - R_{i,X}^*)$ Eq. (SA.9b)

$k = (m - R_{max,X})/(R_{max,X}/R_{i,X}^*) + R_{i,X}^*$ Eq. (SA.9c)

$k = m(R_{i,X}^*/R_{max,X}) - R_{i,X}^* + R_{i,X}^*$ Eq. (SA.9d)

$k/m = (R_{i,X}^*/R_{max,X})$ Eq. (SA.9e)

$R_{max,X} = R_{i,X}^*(m/k)$ Eq. (SA.9f)

Thus, when $R_{max,X} > R_{i,X}^*(m/k)$, then resource Y is more limiting than resource X, and $N_i^* = \max[0, R_{max,Y}/R_{i,Y}^* - 1]$, whereas when $R_{max,X} < R_{i,X}^*(m/k)$, then resource X is more limiting, and $N_i^* = \max[0, R_{max,X}/R_{i,X}^* - 1]$. Note that N_i^* is strictly positive provided that $R_{max,X} > R_{i,X}^*$ when resource X is limiting, and provided that $R_{max,Y} > R_{i,Y}^*$ (i.e. $m - R_{max,X} > k - R_{i,X}^*$, or $m - k + R_{i,X}^* > R_{max,X}$) when resource Y is limiting. We can therefore solve for the expected equilibrium value of N_i^* summed across the entire supply gradient by integrating:

$$\left(\int_{R_{max,X}=R_{i,X}^*}^{R_{i,X}^*\frac{m}{k}} \frac{R_{max,X}}{R_{i,X}^*} - 1 \quad \mathrm{d}R_{max,X}\right) + \left(\int_{R_{max,X}=R_{i,X}^*\frac{m}{k}}^{m-k+R_{i,X}^*} \frac{m-R_{max,X}}{k-R_{i,X}^*} - 1 \quad \mathrm{d}R_{max,X}\right)$$ Eq. (SA.10a)

$$\left(\frac{{R_{max,X}}^2}{2R_{i,X}^*} - R_{max,X}\right)\Bigg|_{R_{max,X}=R_{i,X}^*}^{R_{i,X}^*\frac{m}{k}} + \left(\frac{\left(m-R_{max,X}\right)^2}{2(R_{i,X}^*-k)} - R_{max,X}\right)\Bigg|_{R_{max,X}=R_{i,X}^*\frac{m}{k}}^{m-k+R_{i,X}^*}$$ Eq. (SA.10b)

$$\left(\frac{-R_{i,X}^* m(2k-m)}{2k^2}\right) - \frac{-R_{i,X}^*}{2} + \left(\frac{k-R_{i,X}^*}{2} - m\right) - \left(\frac{-\left(m(2kR_{i,X}^* - mR_{i,X}^* + km\right)}{2k^2}\right)$$ Eq. (SA.10c)

$$\left(\frac{-R_{i,X}^* m(2k-m) + \left(m(2kR_{i,X}^* - mR_{i,X}^* + km\right)}{2k^2}\right) + \left(\frac{k}{2} - m\right)$$ Eq. (SA.10d)

$$\left(\frac{km^2}{2k^2}\right) + \left(\frac{k}{2} - m\right)$$ Eq. (SA.10e)

$$\frac{(k-m)^2}{2k}$$ Eq. (SA.10f)

Note that the solution is independent of $R_{i,j}^*$, meaning that all species share the same monoculture biomass. Moreover, this solution implies that total monoculture abundance across all sites is strictly positive given that some sites exist where $R_{max,Y} > R_{i,Y}^*$ or $R_{max,Y} > R_{i,Y}^*$ (i.e. $m > k$).

Appendix B: General demonstration of asymmetric competition

Appendix B.I: Resource competition model from the main text

In the two resource model described in the main text, for any two species where A is more of a specialist than B, and where A is a better competitor for resource Y while B is a better competitor for resource X, we can show that $\alpha_{A,B} > \alpha_{B,A}$ (i.e. $R_{B,X}^*/R_{A,X}^* > R_{A,Y}^*/R_{B,Y}^*$). To make the tradeoff surface somewhat more general, we can re-write the constraints as $\Sigma_j R_{i,j}^* = k$ for all species i and resources j, where k is a positive number. First, we re-write this inequality as:

$$R_{B,X}^*/(k - R_{A,Y}^*) > R_{A,Y}^*/(k - R_{B,X}^*) \quad \text{Eq. (SB.1a)}$$

$$R_{B,X}^* (k - R_{B,X}^*) > R_{A,Y}^* (k - R_{A,Y}^*) \quad \text{Eq. (SB.1b)}$$

$$k R_{B,X}^* - (R_{B,X}^*)^2 > k R_{A,Y}^* - (R_{A,Y}^*)^2 \quad \text{Eq. (SB.1c)}$$

Note that the function $R_{i,j}^* - (R_{i,j}^*)^2$ is monotonic and increasing over the interval $(0, k/2)$. We can show this with a first derivative test as:

$$d(k R_{i,j}^* - (R_{i,j}^*)^2)/d(R_{i,j}^*) = k - 2(R_{i,j}^*) > 0 \quad \text{Eq. (SB.1f)}$$

$$k/2 > R_{i,j}^* \quad \text{Eq. (SB.1g)}$$

Given the tradeoff $R_{i,X}^* + R_{i,Y}^* - k = 0$, R^* of resources for which species are the "better" competitor (i.e. $R_{A,Y}^*$ and $R_{B,X}^*$) must be $\leq k/2$ (since $R_{i,X}^* > k/2$ requires $R_{i,Y}^* < k/2$). This implies that Eq. (SB.1c) is true if $R_{B,X}^* > R_{A,Y}^*$, which is true by definition because A is more of a specialist than B. Therefore, this also implies that $\alpha_{A,B} > \alpha_{B,A}$.

Appendix B.II: Examples of other models of resource competition

Here, following the same steps as we do for the solution in *Appendix B.I*, we analyze three other types of resource competition for asymmetrical competition between specialists and generalists, using models with dynamical resources following Tilman (1982): ***(1)*** competition for perfectly substitutable resources; ***(2)*** competition for essential resources; and ***(3)*** competition for switching resources. For model ***(1)***, we find the opposite results from those in the main text (i.e. the effect of specialists on generalists is stronger than the effect of generalists on specialists). For model ***(2)***, we find similar results to those presented in *Appendix B.I* (i.e. competitive effect of generalists on specialists is greater than the effect of specialists on generalists). For the third model, we

find similar results to those in the main text under some circumstances, but diverging results under other circumstances.

(1) Competition for two perfectly substitutable resources, *sensu* Tilman (Tilman 1982 p. 271)

In this system, competition of two species A and B for two resources X and Y can be described around equilibrium with the competition coefficients:

$$\alpha_{A,B} = (c_{BY} + c_{BX}\, w_{AX}/w_{AY})/(c_{AY} + c_{AX}\, w_{AX}/w_{AY}) \quad \text{Eq. (SB.2a)}$$

$$\alpha_{B,A} = (c_{AY} + c_{AX}\, w_{BX}/w_{BY})/(c_{BY} + c_{BX}\, w_{BX}/w_{BY}) \quad \text{Eq. (SB.2b)}$$

where c_{ij} is the amount of resource j consumed by species i (per unit biomass, per unit time), and w_{ij} is the amount of biomass accumulated by species i per unit of resource j consumed (similar to the tissue concentration parameters q is *Appendix A*). If we assume optimal foraging, then the ratio of the consumption vectors c_{iX}/c_{iY} for each species should equal the ratio of conversion factors w_{iX}/w_{iY}. Thus, we can re-write Eqs. (SB.2a-b) as:

$$\alpha_{A,B} = (c_{BY} + c_{BX}\, c_{AX}/c_{AY})/(c_{AY} + c_{AX}^2/c_{AY}) \quad \text{Eq. (SB.2c)}$$

$$\alpha_{B,A} = (c_{AY} + c_{AX}\, c_{BX}/c_{BY})/(c_{BY} + c_{BX}^2/c_{BY}) \quad \text{Eq. (SB.2d)}$$

Now, assuming that species consumption vectors fall long a linear tradeoff surface such that $\Sigma_j c_{ij} = k$. We can therefore rewrite Eqs. (SB.2a-b) as:

$$\alpha_{A,B} = (c_{BY} + (k - c_{BY})\, c_{AX}/(k - c_{AX}))/(k - c_{AX} + c_{AX}^2/(k - c_{AX})) \quad \text{Eq. (SB.2e)}$$

$$\alpha_{B,A} = (k - c_{AX} + c_{AX}\, (k - c_{BY})/c_{BY})/(c_{BY} + (k - c_{BY})^2/c_{BY}) \quad \text{Eq. (SB.2f)}$$

Lastly, if A is more of a specialist than B is (i.e. c_{AX} and c_{BX} take on more extreme values along the tradeoff surface than c_{BX} and c_{BY}) then (arbitrarily choosing resource X as the resource for which species A is a better competitor) we can assume:

$$c_{AX} < c_{BX} \quad \text{Eq. (SB.2g)}$$

$$c_{AY} > c_{BY} \quad \text{Eq. (SB.2h)}$$

Note that Eq. (SB.2h) can be re-written as:

$$k - c_{AX} > k - c_{BX} \quad \text{Eq. (SB.2i)}$$

which can be simplified to Eq. (SB.2g) (i.e. Eq. (SB.2g) and Eq. (SB.2h) are equivalent given the tradeoff).

If the criterion $\alpha_{A,B} < \alpha_{B,A}$ is true (i.e. the effect of specialists on generalists is larger than the effect of generalists on specialists), this implies:

$$(c_{BY} + (k - c_{BY})\, c_{AX}/(k - c_{AX}))/(k - c_{AX} + c_{AX}^2/(k - c_{AX})) < \ldots$$
$$(k - c_{AX} + c_{AX}\,(k - c_{BY})/c_{BY})/(c_{BY} + (k - c_{BY})^2/c_{BY}) \quad \text{Eq. (SB.2j)}$$

$$(c_{BY} + (k - c_{BY})\, c_{AX}/(k - c_{AX}))\,(c_{BY} + (k - c_{BY})^2/c_{BY}) < \ldots$$
$$(k - c_{AX} + c_{AX}\,(k - c_{BY})/c_{BY})\,(k - c_{AX} + c_{AX}^2/(k - c_{AX})) \quad \text{Eq. (SB.2k)}$$

$$(c_{BY} + (k - c_{BY})\, c_{AX}/(k - c_{AX}))\,(c_{BY} + (k - c_{BY})^2/c_{BY}) -$$
$$(k - c_{AX} + c_{AX}\,(k - c_{BY})/c_{BY})\,(k - c_{AX} + c_{AX}^2/(k - c_{AX})) < 0 \quad \text{Eq. (SB.2l)}$$

Though it is somewhat complex, note that Eq. (SB.2l) is cubic in c_{AB} and c_{BA}, and therefore equals zero at up to three points. Solving for these yields:

$$c_{AX} = c_{BY} \quad \text{Eq. (SB.2m)}$$

$$c_{AX} = k - c_{BY} \quad \text{Eq. (SB.2n)}$$

$$c_{AX} = (k\, c_{BY})/(2c_{BY} - k) \quad \text{Eq. (SB.2o)}$$

Note that Eq. (SB.2o) is always less than or equal to zero (if $c_{BY} \leq k/2$), or greater than c_{BY} (d/Eq. (SB.2o)/d$c_{BY} = -k^2/(2c_{BY} - k)^2$, which is always less than zero – Eq. (SB.2o) is therefore a monotonic decreasing function of c_{BY}, with a minimum over the range of ($k/2$, k] of Eq. (SB.2o) = c_{BY} when $c_{BY} = k$). We can also ignore the region of c_{AX} that falls between c_{BX} and $k - c_{BX}$, because by definition $c_{AX} < c_{BX}$, and $c_{AX} < k - c_{BY}$. Thus, we can focus solely on the range of c_{AX} values over $(0, c_{BY})$ if $c_{BY} < k/2$, and over $(0, k - c_{BY})$ if $c_{BY} > k/2$.

For both of these intervals, we can substitute $c_{AX} = 0 + \tau$, where τ is an arbitrarily small positive deviation, and re-write Eq. (SB.2o) as:

$$(c_{BY} - (\tau\,(c_{BY} - k))/(k - \tau))*(c_{BY} + (c_{BY} - k)^2/c_{BY}) + \ldots$$
$$(\tau - k + (\tau*(c_{BY} - k))/\, c_{BY})(k - \tau + \tau^2/(k - \tau)) < 0 \quad \text{Eq. (SB.2p)}$$

Because τ is small, this is approximately equal to:

$$(c_{BY})*(c_{BY} + (c_{BY} - k)^2/c_{BY}) - k^2 < 0 \quad \text{Eq. (SB.2q)}$$

$$c_{BY}^2 + (c_{BY} - k)^2 - k^2 < 0 \quad \text{Eq. (SB.2r)}$$

$$c_{BY}^2 - kc_{BY} < 0 \quad \text{Eq. (SB.2s)}$$

which holds for all $c_{BY} < k$. This implies $\alpha_{A,B} < \alpha_{B,A}$.

(2) Competition for two essential resources, *sensu* Tilman (Tilman 1982 p. 199)

Assuming that the species are able to coexist stably at equilibrium, and that species A is a better competitor for resource X while B is a better competitor for resource Y, competition coefficients are:

$\alpha_{A,B} = (c_{BY})/(c_{AY})$ Eq. (SB.3a)

$\alpha_{B,A} = (c_{AX})/(c_{BX})$ Eq. (SB.3b)

Again, assuming a tradeoff where $\Sigma_j c_{ij} = k$, we can re-write this as:

$\alpha_{A,B} = (c_{BY})/(k - c_{AX})$ Eq. (SB.3c)

$\alpha_{B,A} = (c_{AX})/(k - c_{BY})$ Eq. (SB.3d)

Lastly, recall that if A is more of a specialist than B, and that both species forage optimally, then $c_{AX} < c_{BY}$ Eq. (SB.3e)

$c_{AY} > c_{BX}$ Eq. (SB.3f)

Again, note that Eq. (SB.3f) can be re-written as:

$k - c_{AX} > k - c_{BY}$ Eq. (SB.3g)

which can be simplified to Eq. (SB.3e) (i.e. Eq. (SB.3e) and Eq. (SB.3f) are equivalent given the tradeoff).

If the criterion $\alpha_{A,B} > \alpha_{B,A}$ is true (i.e. the effect of generalists on specialists is greater than the effect of specialists on generalists), this implies:

$(c_{BY})/(k - c_{AX}) > (c_{AX})/(k - c_{BY})$ Eq. (SB.3h)

$(c_{BY})(k - c_{BY}) > (c_{AX})(k - c_{AX})$ Eq. (SB.3i)

$k\, c_{BY} - c_{BY}^2 > k\, c_{AX} - c_{AX}^2$ Eq. (SB.3j)

Recall that c_{AX} and c_{BY} (as well as all other consumption vectors) are bounded over the interval $[0, k]$. Eq. (SB.3j) is therefore of the same form as Eq. SB.1c, implying that Eq. (SB.3j)

is true if $c_{BY} > c_{AX}$, which matches our assumption that A is a specialist and B is a generalist. This confirms that $\alpha_{A,B} > \alpha_{B,A}$.

(3) Competition for two switching resources, *sensu* Tilman (Tilman 1982 p. 202)

In this model, competition coefficients are zero if the species can coexist at stable equilibrium. During transient dynamics when both species do have negative effects on one another, however, competition coefficients are given as:

$\alpha_{A,B} = (c_{Bj})/(c_{Aj})$ Eq. (SB.4a)

$\alpha_{B,A} = (c_{Aj})/(c_{Bj})$ Eq. (SB.4b)

where j is a shared limiting resource (either X or Y).

Assuming A is more of a specialist than B (i.e. consumption rates for A are at more extreme positions on the tradeoff surface than they are for B), then we can characterize the relative magnitudes of their consumption rates (again, assuming optimal foraging and arbitrarily choosing resource X as the resource for which species A is a better competitor) as:

$c_{AX} < c_{BX}$ Eq. (SB.4c)

$c_{AY} > c_{BY}$ Eq. (SB.4d)

Next, to determine the relative magnitudes of $\alpha_{A,B}$ and $\alpha_{B,A}$ we can test:

$\alpha_{A,B}$ (?) $\alpha_{B,A}$ Eq. (SB.4e)

$(c_{Bj})/(c_{Aj})$ (?) $(c_{Aj})/(c_{Bj})$ Eq. (SB.4f)

$(c_{Bj})^2$ (?) $(c_{Aj})^2$ Eq. (SB.4g)

For consumption vectors bounded between 0 and k, $(c_{Bj})^2 > (c_{Aj})^2$ is true, trivially, for any case where $c_{Bj} > c_{Aj}$, while $(c_{Bj})^2 < (c_{Aj})^2$ if $c_{Bj} < c_{Aj}$. Note that this implies that $\alpha_{A,B} > \alpha_{B,A}$ only when species are both competing for resource X (i.e. the resource for which the species that is more of a specialist, A, is a better competitor). When both species compete for Y, however, this implies that $\alpha_{A,B} < \alpha_{B,A}$ (i.e. the effect of the specialist on the generalist is stronger than the effect of the generalist on the specialist).

Appendix C: Computational integration methods

To perform the computational integrations described in the main text, we used Euler's method with a fixed time step of 1% of species growth rates, implemented in the C programming language. We used this method because it allowed us to easily stop

integration when the system reduced to an analytically tractable state (i.e. two or fewer species persisting), or when the system neared equilibrium.

For several hundred combinations of sites and species including boundaries between species ranges, we tested our integration method against the more commonly used "`lsoda`" and "`ode45`" algorithms, implemented in the `deSolve` (Soetaert, Petzoldt & Setzer 2010) package in the `R` programming language (R. Development Core Team, 2016) using several different starting population sizes. All methods and cases produced identical results.

The model we use is very similar to that of May and MacArthur (1972), for which any stable multi-species community is globally attracting given nonzero species abundances and fixed environmental conditions. Though our model differs slightly in that $\alpha_{A,B} \neq \alpha_{B,A}$, which makes an analytical proof of global stability in our system difficult, in numerical simulations of the system for communities of between 2 and 50 species, we find results consistent with global stability in all cases tested (i.e. the same equilibrium is approached from any starting point with nonzero species abundances). Thus, we are reasonably confident that the computational solutions we generated are comparable to the analytical solutions for the simpler two resource model.

Tables:

Species Name	R_N^* mg kg-1	E026	E055	E111	E120	E123	Total
Achillea millefolium L.	0.131	0	0	1	2	2	5
Agrostis scabra Willd.	0.156	27	0	2	0	0	29
Ambrosia artemisiifolia L.	0.149	0	0	4	0	0	4
Andropogon gerardii Vitman	0.066	7	3	4	3	2	19
Anemone cylindrica A.Gray	0.146	0	0	4	0	2	6
Asclepias syriaca L.	0.255	0	0	4	0	0	4
Asclepias tuberosa L.	0.202	0	0	4	2	1	7
Asclepias verticillata L.	0.144	0	0	1	0	0	1
Calamovilfa longifolia (Hook.) Scribn.	0.14	0	0	4	0	0	4
Dalea purpurea[1] Vent.	0.399	0	3	4	3	1	11
Echinacea serotina[2] (Nutt.) DC.	0.256	0	0	3	0	0	3
Elymus repens[3] (L.) Gould	0.135	21	0	4	0	0	25
Koeleria macrantha[4] (Ledeb.) Schult.	0.142	0	0	4	2	0	6
Lespedeza capitata Michx.	0.25	0	3	4	4	1	12
Liatris aspera Michx.	0.143	0	0	2	3	0	5
Penstemon grandiflorus Nutt.	0.115	0	0	4	0	0	4
Poa pratensis L.	0.103	5	0	3	0	0	8
Schizachyrium scoparium (Michx.) Nash	0.047	8	3	4	3	2	20
Solidago nemoralis Aiton	0.097	0	0	4	0	0	4
Solidago rigida L.	0.083	0	0	4	3	0	7
Sorghastrum nutans (L.) Nash	0.091	0	0	4	4	1	9
Stipa spartea Trin.	0.21	0	0	4	0	0	4
Symphyotrichum oolentangiense[5] (Riddell) G.L.Nesom	0.223	0	0	2	0	0	2
Symphyotrichum ericoides[6] (L.) G.L.Nesom	0.151	0	0	3	0	0	3

Common synonyms: [1]*Petalostemum purpureum*; [2]*Rudbeckia serotina*; [3]*Agropyron repens*; [4]*Koeleria cristata*; [5]*Aster azureus*; [6]*Aster ericoides*

Table S1: R^* for soil nitrate, R_N^*, for species measured in experimental monocultures at Cedar Creek. "Measured Soil Nitrate, mg kg-1" shows R_N^* for found in the nitrogen addition experiment analyzed in Fig. 6. Numbers and labeled columns describe the sample size and data source for each species. Complete data and methods are available as follows: E026, Wedin and Tilman (1993); E055, Dybzinski and Tilman (2007); E111, Craine et al. (2002); E120, Tilman et al. (1997); E123, Tilman et al. (1996).

Appendices for Chapter Four:

Tradeoff-Based Mechanisms Predict Coexistence, Productivity, and Species Abundances in Grassland Plant Communities

Appendix I. Data Collection

As described in the main text, we used data from monocultures in seven experiments: Three competition experiments (E026, E055, E070), three diversity experiments (E120, E123, E249), and one monoculture garden (E111). We also compared regression and mechanism-based model predictions to multi-species mixtures in E120. Full methods for E026 are described in Wedin and Tilman (1993); E055 and E070 (which was a subset of E055) are described in Dybzinski and Tilman (2007); E111 is described in Craine (2002); E123 is described in Tilman *et al.* (1996); E120 is described in Tilman *et al.* (1997); and E249 (which is nested within E120) is described in Whittington *et al*. (2013) and Cowles *et al.* (2016).

Describing each experiment briefly:

E026: Established in 1986. Intended as a test of soil nutrient gradients on competitive hierarchies: we used only data from monocultures grown in nitrogen poor soils. Plots were 0.75m by 0.75m. Vegetation was sampled using 0.1m by 0.4m clip strips. We include data from sampling done in 1988, 1990, and 1992.

E055: Established in 1988. Intended as a test of soil nutrient gradients on competitive hierarchies: we used only data from monocultures grown in nitrogen poor soils. Plots were 1.1m by 1.1m. Vegetation was sampled using 0.1m by 0.5m or 0.1m by 0.1m clip strips. We include data sampling done in 1992, 1996, and 1999. Annual burning treatments to control litter build-up (conducted in the spring before biomass greening) began in 1991.

E070: Same establishment and sampling data as for E055 (though clip strips were always 0.5m long). We use data collected in 1998.

E111: Established in 1992. Included only monocultures grown in un-amended soils. Plots were either 1.5m by 2.4m, or 1.5m by 1.2m. Vegetation was sampled using 0.1m by 2.3m

clip strips in the larger plots, and two 0.1m by 1.2m clip strips in smaller plots. We use data collected in 1997.

E120: Established in 1994, but still ongoing. Experiment includes plots planted with 1, 2, 4, 8, and 16 species. We used data from both monocultures and multi-species mixtures in this experiment. Plots are 9m by 9m. Vegetation is sampled using four 0.1m by 1.5m clip strips per plot. We include sampling conducted annually between 1997 and 2014. Plots are burned annually in the spring to prevent litter build-up. Two grasses, *Agropyron smithii* and *Elymus canadensis*, never established in monocultures or multi-species mixtures in E120, and two oak species, *Quercus ellipsoidalis* and *Q. macrocarpa*, were removed or competitively excluded from most plots. For our analyses, we therefore treated all plots as though *A. smithii* and *E. canadensis* had not been planted, and excluded all plots where oaks were present from analyses. One species, *Amorpha canescens*, is present only in the 1 and 16-species plots in this experiment.

E123: Established in 1994. Includes plots planted with various numbers of species, but we use only data from monocultures. Plots were 3m by 3m. Vegetation was sampled using 0.1m by 3m clip strips. We use data collected in 1997.

E249: Built into a subset of the plots in E120. Established in 2008, but still ongoing. Includes experimental warming treatments, but we use data only from un-manipulated monocultures. Plots are divided by treatment into 2.5m by 3m subplots. Vegetation is sampled using two 0.1m by 1.5m clip strips. We use data collected in 2011 and 2012. Plots are burned annually in the spring to prevent litter build-up.

All data are available through the LTER data portal (lternet.edu/sites/cdr), or at the Cedar Creek website (cbs.umn.edu/explore/field-stations/cedarcreek). See Appendix *XII* for directions on how to access the source code used to implement data cleaning and standardization. To assure comparability of data across these experiments we use only data:

1. *From plants grown on nitrogen-limited soils.* E026, E055, and E077 included soil nitrogen gradients, and E026 included a fertilization treatment. We excluded the

fertilized treatment, and all soil nitrogen treatments that were more than one standard deviation above the mean 1994 total soil nitrogen concentration for plots in E120.

2. *With low litter.* E026 was not burned annually, which lead to litter accumulation. We removed all observations where litter was more than one standard deviation above the mean litter abundance in E120.
3. *For well-established monocultures.* Some monocultures did not establish well. We excluded any observations with aboveground biomass below 10 g m^{-2}. Similarly, *Poa pratensis* did not establish well in its planted monocultures in E120, but did do well in more diverse plots (likely because of limited water availability in monocultures). Because of this, we used only monoculture measurements from experiments other than E120 for this species.
4. *With complete sampling.* We excluded any years that did not include measurements of a majority of species planted in the experiment, to reduce bias from random variation across years.
5. *For herbaceous perennials.* We included only species that fell into one of four functional groups: C3 and C4 grasses, non-legume forbs (flowers), and legumes. We excluded woody species, because these were usually not allowed to mature over the course of experiments, and sedges, because we did not have sufficient species coverage for them.
6. *For mature, maintained plants.* We only included data collected three or more years after monocultures were planted to avoid establishment effects. Similarly, we excluded any measurements from plots that were taken after weeding treatments had ceased.

Appendix II. Trait measurements

For each of the traits that we report, we estimated mean values and within-species trait variation using linear mixed-effects models implemented through the `lmer` function in the `lme4` package (Bates *et al.* 2015) in the `R` programming language (R. Development Core Team, 2016). To account for repeated measurements and spatial pseudo-replication, we fit each model with a random intercept with respect to subplot (if present), nested within plot, nested within experiment, following the form: trait ~ species

+ (1|experiment/plot/subplot). We used the fixed effect parameter estimates and standard errors from these fitted models for trait estimates. To meet distributional assumptions, we log-transformed R^* and B^*_{mono}, and logit-transformed q prior to all analyses.

In order to assure that mean estimates for monocultures from our fitted models matched observations from monocultures in E120, we centred estimates for $B_i^*_{mono}$ for all species in E120 to match the mean monoculture biomass observed in experimental plots in E120 (except for *P. pratensis* as outlined above). We did not centre means prior to fitting the tradeoff surface. For the model with snapped traits, we adjusted q_i after centring means in order to maintain the original $q_i B_i^*_{mono}$ (and therefore coexistence criteria) predicted from the tradeoff surface.

To account for variability in initial soil fertility (S) among plots in E120, we multiplied $B_i^*_{mono}$ by C_k/C, where C_k is the pre-treatment (in 1994) total carbon content of soil in plot k, and C is the mean pre-treatment total carbon content across all plots in E120. We use this metric for fertility because soil carbon concentration is more closely correlated with nitrogen mineralization rates at our site than is soil nitrogen concentration (or any other easily measured soil characteristics) (Fornara & Tilman 2008). This ratio was meant to represent S_k/S, which determines changes in biomass due to soil fertility following Eq. (2) in the main text: $B_{i,j}^*_{mono} = (1 - m_i/c_i)S_{i,k}/q_i$, which implies $B_{i,j}^*_{mono}/B_i^*_{mono} = S_{i,k}/S$.

Appendix III. Observed pairwise associations in E120

To demonstrate the pairwise relationships among species observed in multi-species communities in E120, we fit linear regressions relating the biomass of inferior competitors (based on snapped R^* values) to the biomass of all superior competitors in each plot. For all regressions, we used log-transformed biomass + 0.01 gm^{-2} in order to account for zero observations, and included pre-treatment total soil carbon, year, and their interaction as covariates. Given an inferior competitor species i and a superior competitor species *1, 2, 3...*, this yielded a model of the form:

$$\log_{10}(B_i) = \beta_0 + \beta_1 \log_{10}(\%\text{ soil C}) + \beta_2 \text{year} + \beta_3 \log_{10}(\%\text{ soil C})*\text{year} + \beta_4 \log_{10}(B_1) + \beta_5 \log_{10}(B_2) + \beta_6 \log_{10}(B_3) + ... + \varepsilon$$

where β_i are fitted coefficients, and ε is unexplained residual error.

To estimate the statistical effect of each superior competitor on inferior competitors, we fit a model with all species but the focal superior competitor, and then used backwards selection based on regression AIC to remove terms from the model (e.g. to estimate the effect of species 1 on species 4, we fit a model with % soil C, year, and species 2 and 3, and then ran backwards selection). This produced a "reduced" model of covariates, including % soil C, year, and other potential competitors. Finally, we fit a regression comparing residual variance from this reduced model to residual variance comparing the abundance of the focal superior competitor to the full suite of covariates. This is analogous to building an added variable plot or conducting a type III sum of squares regression, and provided an estimate of the marginal statistical effect of a single superior competitor on an inferior competitor, after controlling for the effects of potential covariates.

In most cases, superior competitors had either negative, or non-significant statistical effects on inferior competitors (Fig. S1). This included negative associations between legume abundance and the abundance of several grass species (with grasses predicted to be competitively superior to the legumes). The only consistent positive statistical effects were of other species on the abundance of *P. pratensis*, which is probably a symptom of failed establishment in low-diversity mixtures.

Appendix IV. Trait-based regressions

We fit linear regressions estimating species abundances in the multi-species mixtures using the three traits identified in the main text (i.e. R^*, q, and B^*_{mono}). Regressions included each trait individually as predictor variables, all three traits, or all two- or three-way interactions among these traits. As described in the main text, we used log transformations for biomass all all traits but q, for which we used a logit transformation. Models also included one or two covariates, which were interacted with all other terms in the model. First, all models included percent soil carbon measured at the start of the E120 experiment (i.e. in 1994), as described in Appendix *II*. We included this term in order to account for its role in mechanistic model predictions. Second, we fit

all models with and without planted richness as a covariate. This was to test for changes in the association between traits and estimates of species abundances relating to mixture diversity.

All models were fit using the `lm` function in `R`. Regressions were fitted to mean species abundance in each plot (i.e. averaged across all observations between 2001 and 2014). As such, regressions followed the form: `lm(abundance ~ (traits and interactions)*(covariates))`. We also used these regressions to estimate total plot biomass, by summing together predictions of species abundances in each plot. We did this, rather than fit a regression to total observed plot-level biomass, because this allowed us to use species-level trait data in a relatively simple regression framework (as opposed to having to include information about multiple species traits per plot). To assess model fit, we used the `lmodel2` function (Legendre 2014) to conduct ranged major axis regressions comparing observations to regression predictions. This provided estimates of model fits that were directly comparable to those we generated for the mechanism-based models. To account for the number of fitted parameters in each regression, we calculated adjusted R^2 for the regression models, as $R^2_{adj} = 1 - [(1 - R^2)(n - 1)/(n - k - 1)]$, where n was the number of observations, and k was the number of fitted parameters. However, this had minimal effect, and differences between R^2 and R^2_{adj} were never greater than 0.015.

For each regression, we also calculated the partial R^2 for each predictor variable, describing the relative contribution of each variable to total model fit. To calculate partial R^2, we refit each model after removing each term sequentially, and calculated the change in fit as partial $R^2 = (SSE_{reduced} - SSE_{full})/SSE_{reduced}$, where SSE_{full} and $SSE_{reduced}$ describe the summed square error of the full and reduced models, respectively.

Appendix V. Major axis regression

To fit tradeoff surfaces and compare observations and predictions, we used major axis regression (type II regression models), which minimizes the squared Euclidian distance between observed points and the fitted surface (i.e. the point where the tangent connecting the regression line and the observation meet the regression plane, or formally,

the projection of the observed points onto the regression surface). Given an n-dimensional regression fit $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n + \beta_0 = 0$, for observed values x_1, x_2, x_3, …, x_n, and fitted parameters β_1, β_2, β_3, …, β_n, the "snapped" coordinates are found by solving $x_{i,snap} = x_i - \beta_i k$, where $k = (\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ...\beta_n x_n + \beta_0)/(\beta_1^2 + \beta_2^2 + \beta_3^2 + \ldots + \beta_n^2)$. We use this method because standard least squares regression assumes that there is no error in the predictor variables, whereas we have no a priori reason for assuming error is absent in any of the variables we consider. We implement this algorithm in the `nondirfit` function, available in the `get_filtered_estimate_functions.R` script (see Appendix *XII.* for instructions on accessing source code).

We fit the regression to the mean observed trait values for each species. To test whether the relationships among traits matched those predicted by our model (i.e. decreases in R^* must correspond to decreases in q or B^*_{mono}), we estimated confidence intervals and covariance among regression parameters using nonparametric bootstrapping. For each of 20,000 iterations, we resampled species traits, with replacement, from the pool of all species, and re-calculated the model parameters for each iteration. We then tested the strength and direction of relationships among traits to determine their significance. This also allowed us to test the significance of a three-way negative relationship among traits, as predicted by our resource competition model.

Appendix VI. Model equilibria and stability

The model we present in the main text can be written in terms that match the classical Lotka-Volterra competition equations, and a general proof of global stability for our system is available in Theorem 3.3.1 of Takeuchi (1996), (p. 36). A briefer proof is available in Takeuchi *et al.* (1978) in Theorem 4, though this 1978 journal article may prove easier to access than the 1996 book. Here, we *(i)* include a somewhat simpler and relatively robust proof of global stability for this model using linearized stability analysis (i.e. calculation of the dominant eigenvalue at equilibrium), and *(ii)* show how the same solution can be derived using Chesson's pairwise niche overlap and fitness difference framework (Chesson 1990, 2000, 2013; Letten, Ke & Fukami 2017).

VI.i. Linearized stability analysis

First, we begin with a classical Lotka-Volterra competition model of the form:

$$dN_i/dt = r_iN_i(1 - (N_i + \alpha_{ji}N_j)/K_i) \quad (SVI.i.1)$$

Starting with Eq. (1) from the main text, we can define biomass dynamics with species *i* competitively inferior to species *j* as:

$$dB_i/dt = c_iB_i(1 - (q_iB_i + q_jB_j)/S) - m_iB_i \quad (SVI.i.2)$$

Redistributing terms yields:

$$\begin{aligned} dB_i/dt &= c_iB_i(1 - (q_iB_i + q_jB_j)/S) - m_iB_i \quad (SVI.i.3)\\ &= (c_i - m_i)B_i - c_i(q_iB_i + q_jB_j)/S)(c_i - m_i)/(c_i - m_i)\\ &= (c_i - m_i)B_i(1 - c_i(q_iB_i + q_jB_j)/((c_i - m_i)S))\\ &= (c_i - m_i)B_i(1 - c_i(q_iB_i + q_jB_j)/((c_i - m_i)S))\\ &= (c_i - m_i)B_i(1 - (q_iB_i + (q_i/q_i)q_jB_j)/((1 - m_i/c_i)S))\\ &= (c_i - m_i)B_i(1 - (B_i + (q_j/q_i)B_j)/((1 - m_i/c_i)S/q_i)) \end{aligned}$$

Substituting $N_i \equiv B_i$, and $N_j \equiv B_j$, and $K_i \equiv (1 - m_i/c_i)S/q_i \equiv B_i{}^*_{mono}$ in (S6) yields:

$$dN_i/dt = (c_i - m_i)N_i(1 - (N_i + (q_j/q_i)N_j)/(K_i) \quad (SVI.i.4)$$

From equation (S*VI.i.*1), this implies that $r_i = (c_i - m_i)$, and $\alpha_{ij} = q_j/q_i$. Note that competition in the reverse direction (i.e. the effect of inferior competitors on superior competitors) is simply $\alpha_{ji} = 0$. Because of this, equilibrium conditions can be derived sequentially, starting with the best competitor (for which $B_{i'}{}^* = K_{i'}$), and proceeding through sequentially worse competitors based on α_{ij} and $B_j{}^*$.

Given $\alpha_{ij} < 0$ and $\alpha_{ji} = 0$, Eq. (S*VI.i.*1) shows that $\partial(dN_i/dt)/\partial N_j < 0$ for all populations where $N_i > 0$ and $N_j > 0$. To evaluate $\partial(dN_i/dt)/\partial N_i$ at equilibrium, we define the effect of all superior competitors on species *i* as $\Sigma_{j<i}\,(q_j/q_i\,N_j{}^*) \equiv \kappa_i$. From Eq. (3) in the main text, this implies that for the non-trivial equilibrium, $N_i{}^* = K_i - \kappa_i$. Substituting this into (S*VI.i.*4) at equilibrium yields:

$$dN_i/dt|[Ni = N_i{}^*] = r_i(K_i - \kappa_i)(1 - ((K_i - \kappa_i) + \kappa_i)/(K_i) = 0 \quad (SVI.i.5)$$

Given a small change in population size of species 1, η_i, this can be re-written as:

$dN_i/dt|[Ni = N_i^* + \eta_i] = r_i(K_i + \eta_i - \kappa_i)(1 - ((K_i + \eta_i - \kappa_i) + \kappa_i)/(K_i)$ (S*VI.i*.6)

Calculating the partial derivative of (S*VI.i*.6) with respect to η_i yields:

$\partial(dN_i/dt)/\partial\eta_i = -r_i((K_i + \eta_i)/K_i - 1) - r_i(K_i - \kappa_i + \eta_i)/K_i$ (S*VI.i*.7)

Given $\eta_i > 0$, the first term is always negative, because $(K_i + \eta_i)/K_i > 1$. The second term is always negative for any $K_i < \kappa_i$ (i.e. $N_i^* > 0$). Because η_i describes changes in N_i around equilibrium, this implies that $\partial(dN_i/dt)/\partial N_i\ |[N_i = N_i^*] < 0$.

Following from this, the Jacobian matrix describing interspecific interactions for any number of competing species in this model at equilibrium is a triangular matrix with purely negative non-zero elements. Because the diagonal elements of a triangular matrix are also its eigenvalues, this implies that all eigenvalues for the system are also negative. Thus, the equilibrium is locally stable, given any mixture of species with positive abundances at equilibrium. As described in Eq. (3) and Eq. (5b) in the main text, this requires $K_i > \Sigma_{j<i}\ (q_j/q_i)K_j$, for all species with $R_j^* < R_i^*$ (i.e. any feasible equilibrium is also stable).

This locally stable equilibrium can also be shown to be globally stable. Because dynamics of all superior competitors are independent from dynamics of inferior competitors, we can address dynamics of inferior competitor *i* given equilibrium abundances of all superior competitors in the community as:

$dN_i/dt = r_iN_i(1 - (N_i + \kappa_i)/(K_i)$ (S*VI.i*.8)

This system has only two possible equilibria: $N_i^* = K_i - \kappa_i$, and the trivial equilibrium $N_i^* = 0$. Following the procedure in (S*VI.i*.5)-(S*VI.i*.7) to solve for $\partial(dN_i/dt)/\partial N_i$ at $N_i = 0$ yields:

$\partial(dN_i/dt)/\partial\eta_i = -r_i(\eta_i/K_i - 1) + r_i(\kappa_i + \eta_i)/K_i$ (S*VI.i*.9)

Given $K_i > \kappa_i$, this is strictly positive for small deviations (i.e. $\eta_i < K_i$ and $\eta_i < \kappa_i$), implying that the diagonal elements of the (triangular) Jacobian matrix are positive, and that its eigenvalues are therefore also positive and that the trivial equilibrium is unstable. The equilibrium $N_i^* = K_i - \kappa_i$ is therefore the only stable equilibrium for species *i*. Finally, from (S*VI.i*.8), $dN_i/dt > 0$ for all $0 < N_i < K_i - \kappa_i$, and $0 > dN_i/dt$ for all $K_i - \kappa_i < N_i$. This implies that the system will approach the stable positive equilibrium from any feasible starting points, and is therefore globally stable at $N_i^* = K_i - \kappa_i$.

VI.ii. Chesson's pairwise method

Given a system with two competing species of the form:

$$dN_1/dt = r_1N_1(1 - a_{11}N_1 - a_{12}N_2) \quad \text{(S}VI.ii\text{.1a)}$$

$$dN_2/dt = r_2N_2(1 - a_{22}N_2 - a_{21}N_1) \quad \text{(S}VI.ii\text{.1b)}$$

Chesson defines niche overlap as:

$$\rho = [(a_{12}a_{21})/(a_{11}a_{22})]^{1/2} \quad \text{(S}VI.ii\text{.2a)}$$

and absolute fitness difference of species 2 relative to species 1 as:

$$f_2/f_1 = [(a_{11}a_{12})/(a_{22}a_{21})]^{1/2} \quad \text{(S}VI.ii\text{.2b)}$$

Derivations and discussions of these terms are available in several works by Chesson (1990, 2000, 2013), and in Letten *et al.* (2017). In this framework, stable coexistence is achieved when niche overlap is less than absolute fitness differences, following:

$$\rho < f_2/f_1 < 1/\rho \quad \text{(S}VI.ii\text{.3)}$$

Following the re-parameterization of the model outlined in Eq. (*VI.i*.1), we can write the interaction terms in Eqs. (S*VI.ii*.1a-b) as:

$$a_{ii} = 1/K_i \quad \text{(S}VI.ii\text{.4a)}$$

$$a_{ij} = (q_j/q_i)(1/K_i) \quad \text{if } R_j^* < R_i^* \quad \text{(S}VI.ii\text{.4b)}$$

$$= 0 \quad \text{otherwise}$$

Because a_{ij} is always zero for any pair of species where $R_i^* < R_j^*$, pairwise interactions between any two species with $R_1^* < R_2^*$ results in $a_{21} > 0$ and $a_{12} = 0$ (i.e. species 2 is the inferior competitor). Substituting these into Eq. (S*VI.ii*.3) and taking the limit as a_{12} approaches zero, we find that the species can stably coexist provided that:

$$[(a_{21})/(a_{11}a_{22})]^{1/2} < [(a_{11})/(a_{22}a_{21})]^{1/2} \quad \text{(S}VI.ii\text{.5)}$$

Note that the upper limit described in the third term of the inequality can now be omitted, because $1/\rho$ necessarily tends towards infinity as a_{12} approaches zero, while the other two terms tend towards zero. Substituting Eqs. (S*VI.ii*.4a-b) into Eqs. (S*VI.ii*.5), we can rewrite the inequality as:

$$[\,(q_1/q_2)K_1]^{1/2} < [(K_2)^2/((q_1/q_2)K_1]^{1/2} \quad \text{(S}VI.ii\text{.6)}$$

From this, we can also rewrite the terms for niche overlap and fitness differences in Eqs. (*VI.ii*.2a-b) as:

$$\rho = [(q_1/q_2)K_1]^{1/2} \quad \text{(S}VI.ii.\text{7a)}$$

$$f_2/f_1 = [(K_2)^2/((q_1/q_2)K_1)]^{1/2} \quad \text{(S}VI.ii.\text{7b)}$$

Note that by multiplying both side of Eq. (*VI.ii*.6) by $[(q_1/q_2)K_1]^{1/2}$, we can further simplify the inequality to:

$$(q_1/q_2)K_1 < K_2 \quad \text{(S}VI.ii.\text{8)}$$

Lastly, recall that this model has a strict competitive hierarchy, in which species with higher values of R^* have no effect on the growth of species with lower R^*. As such, we can calculate equilibrium population abundances sequentially, starting with the best competitor, in order to determine the abundance of each species at equilibrium. This also means that we can abstract the pairwise solution in Eq. (S*VI.ii*.8) to systems with many competing species, by substituting species 1 with the summed effects of all superior competitors on species 2. This yields a general criterion for stable coexistence:

$$\Sigma_{j<i}\,(q_j/q_i)K_j < K_i \quad \text{(S}VI.ii.\text{9)}$$

for all species j that are superior competitors to species i (i.e. all species with $R_j^* < R_i^*$). Note that this inequality is identical to Eq. 5b in the main text, and to the solution derived in *Appendix IX.i*.

Appendix VII. Making model predictions

To predict species-level biomass for the polyculture mixtures planted in E120, we parameterized Eq. (3) in the main text using raw species traits or snapped trait estimates, and solved it sequentially from the best competitor to the worst competitor. This provided an analytically tractable prediction of mean species-level biomass at equilibrium.

Parameterizing the model with mean trait values (i.e. no intraspecific trait variation) under-predicted the number of coexisting species across all multi-species communities (Fig. S4A). To account for within-species trait variation, we calculated observed sample variance among monoculture replicates of each species, and included this in the model either as variability around raw trait values, or around snapped points on the tradeoff surface. This substantially improved predictions of coexistence for both types of models (Fig. S4B). We therefore used this method to account for intraspecific trait variability for all subsequent predictions.

To account for within-species variability in species traits, we sampled 20,000 separate estimates of species traits for each planted mixture, based on the standard error in traits as calculated and reported in Appendix *II*, and either the mean raw traits for each species, or the snapped traits from the tradeoff surface. We then used these 20,000 trait estimates to generate predictions of species-level aboveground biomass, plot-level biomass, and community stability, and calculated mean and standard deviation from the resulting distributions.

We compared our predictions to species-level aboveground biomass observations in E120 measured between 2001 and 2014. To calculate mean biomass across repeated observations or predictions, we log-transformed the values, and used a hurdle model to account for instances of zero biomass. Thus, the mean over "n" potential biomass values was calculated as $B_i = z_i \exp(\Sigma_{Bi*>0} \log(B_i*)/(n\ z_i))$, where z_i is the proportion of observations or predictions where $B_i > 0$. Variability was calculated solely from non-zero, log-transformed biomass values.

Appendix VIII. Assessing model fit

We used the two-sample Wilcoxon test to determine significance of differences between observed and estimated species richness in Fig. 4A-B, differences in error between models parameterized with raw and snapped traits in Fig. 4C-D, and differences in original and augmented model fit in Fig. 5. For all tests, we compared quantities paired by plot (i.e. observed vs. estimated richness in each plot, error from raw vs. snapped traits in each plot, etc.).

In most cases, we calculated R^2 from the difference between observed and predicted values as:

$$1 - \Sigma_i(observed_i - predicted_i)^2/\Sigma_i(observed_i - \text{mean}[observed])^2 \quad \text{(S1)}$$

The exceptions to this are the results in Table 1, Fig. 4A-D, and in Fig. S5, where we show R^2 for the fitted regression line comparing observed and predicted values (i.e. total fraction of variation in observed values explained by predicted values).

For regressions between pairs of variables (Fig. 4 E-H; Fig 4 A-E), we used the `lmodel2` (Legendre 2014) function in `R` to fit ranged major axis regression, and to

calculate p-values. Because this function only accepts bivariate data, we implemented our own function in `R` (`nondirfit`) to handle higher multivariate regressions (e.g. the tradeoff surface among three traits), with details described in Appendix *V*. This function uses a nonlinear optimizer to minimize the distance between standardized raw and snapped trait values (i.e. (x_i-mean(x))/sd(x)), and perfectly matches outputs from the `lmodel2` function for bivariate ranged major axis regression.

To estimate significance of estimates from the major axis regressions fit using the `nondirfit` function, we used a simple bootstrapping routine. Given *n* observations, we sampled with replacement from the total pool of observations *n* times, and fit the regression using the sampled dataset. We then repeated this procedure for 20,000 iterations to generate a distribution of estimates for all terms in the regression. For the tradeoff surface, we used the multivariate distribution of all slope parameters to calculate the proportion of iterations for which there was a negative relationship among all three traits included in the tradeoff.

Appendix IX. Simulating communities from the tradeoff surface

We sampled simulated communities from across the fitted tradeoff surface using the covariance relationship among $\log_{10}$-transformed R^* and B^*_{mono}, and logit-transformed q, using the `mvtnorm` package in `R` (Genz & Bretz 2009; Genz *et al.* 2015). For each of 20,000 iterations, we sampled species from the tradeoff surface, and randomly assembled them into mixtures of species matching the planted richness in plots in E120 to predict total community biomass, and changes in species-level biomass relative to monocultures as a function of other competitors in the community.

To characterize communities in E120, we constrained sampling from this surface, because the species used to assemble E120 were not sampled randomly from across trait space (and because there were not enough species in E120 to directly parameterize the multivariate normal distribution only from species in E120). Rather, the experiment included four species from each of the following functional groups: C3 grass, C4 grass, non-legume forbs, and legumes. These differ substantially in many traits.

We constrained our sampling routine in two ways. First, for each iteration, we constrained samples such that the total plant community matched the functional group distribution found in E120 – 2 C3 grasses (4 species, minus *A. smithii* and *E. canadensis*, which did not germinate), 4 C4 grasses, 4 non-legume forbs, and 3 or 4 legumes (4 species, minus *A. canadensis* in 2, 4, and 8 species mixtures, where it was not planted). To do this, we only included random samples from the tradeoff surface where traits fell within the observed ranges of these functional groups based on the 35 species used to fit the surface. Second, within each iteration, we randomly assembled the simulated species into multi-species communities of varying richness, where probability of drawing a species from a particular functional group matched the average planted proportion of that functional group in E120. Thus, while our sampling routine ensured that simulated species traits would be similar to those of the functional groups planted in E120, it did not require any specific knowledge of the planted species or functional group composition of any individual plot.

Appendix X. Constraints from the empirical tradeoff surface

Because the empirical tradeoff that we discuss in the main text imposes strict relationships among R^*, q, and B^*_{mono}, it is possible that many of our findings, such as increased coexistence and transgressive over-yielding, are trivial outcomes resulting from the negative correlation among traits. However, these results turn out not to be trivial because they rely not only on the *shape* of the tradeoff surface, but also the *distribution of traits* across the surface. In this section, we demonstrate that neither coexistence nor transgressive over-yielding are assured for species with traits that fall along the empirical tradeoff surface (or any tradeoff surface that imposes simple linear constraints on transformed parameter values, as outlined in Appendix *V*).

X.i. Coexistence on the tradeoff surface

As noted in the text and in Appendix *VI*, coexistence between two species in our resource competition model with $R_i^* < R_j^*$ requires that $q_i B_i^*_{mono} < q_j B_j^*_{mono}$. This implies

that $dqB^*_{mono}/d(-R^*) < 0$. Given the fitted tradeoff surface $\beta_1 \log_{10}(B^*_{mono}) + \beta_2 \text{logit}(q) + \beta_3(-\log_{10}(R^*)) + \beta_0 = 0$, we can locally approximate this coexistence criterion as:

$$d[\text{logit}(q)\ \log_{10}(B^*_{mono})]/d[-\log_{10}(R^*)] = -\beta_3(\beta_1 \log_{10}(B^*_{mono}) + \beta_2 \text{logit}(q))/(\beta_1\beta_2) \quad (S\mathit{X}.i.1)$$

because $\log_{10}(R^*)$ scales monotonically with R^*, and $\text{logit}(q)\ \log_{10}(B^*_{mono})$ scales monotonically with qB^*_{mono}. To assure coexistence between all potential pairs of species on the tradeoff surface, this therefore requires:

$$\log_{10}(B^*_{mono}) > -\beta_2/\beta_1 * \text{logit}(q) \quad (S\mathit{X}.i.2)$$

However, the fitted surface need not satisfy the requirement in equation (S*X*.*i*.2), as $\log_{10}(B^*_{mono})$ and $\text{logit}(q)$ vary in an inverse, but otherwise unconstrained, relationship for any given value of R^*. This implies that any of the species that we observed in our system could potentially be competitively excluded by some other species, even if both fall somewhere along our fitted tradeoff surface. This is exemplified in Fig. 2E in the main text, where many species can be competitively excluded by other species even after snapping traits to the empirically observed tradeoff surface.

X.ii. Over-yielding on the tradeoff surface

In our model (which reduces to Lotka-Volterra competition), over-yielding (i.e. greater mean productivity in multi-species mixtures than in the average monoculture) occurs for any species that stably coexist (Loreau 2004). Transgressive over-yielding (i.e. greater mean productivity in multi-species mixtures than in any constituent monoculture) occurs among coexisting species when inferior competitors have higher tissue nitrogen concentration than do superior competitors (i.e. $q_i > q_j$) (Loreau 2004). Under these circumstances, superior competitors produce more biomass per unit nitrogen uptake than inferior competitors do, but inferior competitors nevertheless increases total community biomass when present because they are able to access nitrogen that superior competitors cannot. We can demonstrate this mathematically following Eq. (3) in the main text, which shows that the equilibrium biomass of species i in mixture, B_i^*, is defined as:

$$B_i^* = B_i^*_{mono} - \Sigma_{j<i}\ (q_j/q_i)\ B_j^* \quad (S\mathit{X}.ii.1)$$

where species are arranged in terms of competitive ability, such that $j < i$ implies that j is a superior competitor. Note that when $q_i > q_j$, this implies that the negative effect of

species *j* on the biomass of species *i* is less than the total biomass of species *j*, meaning that species *i* and *j* growing together in mixture will produce more biomass than either growing alone.

As with coexistence, transgressive over-yielding is not imposed by the shape of the empirical tradeoff or by the assumptions in our model, and depends on both the specific shape of the tradeoff surface, and the distribution of traits across it. Nevertheless, transgressive over-yielding (i.e. $q_j > q_i$) is realized for almost all communities that we tested (see Fig. S2B,E). This result also accords with experimental evidence from Cedar Creek that shows productivity is maximized in mixtures that include both species that are efficient nitrogen utilizers (e.g. C4 grasses) and effective at accessing multiple pools of nitrogen (e.g forbs and legumes) (Tilman *et al.* 1997).

Appendix XI. Augmenting model with additional factors

In the main text, we discuss three specific factors omitted from in our model which lead to significant improvements in model fit when included – changes in competitive hierarchy with soil fertility, legume competition for resources other than nitrogen, and seasonal changes in species traits. Here, we discuss specific details for how we included these additional strategies. Results from analyses are presented in Fig. S5.

To test for the effects of changes in competitive hierarchy with soil fertility, we utilized data from monocultures of the two species with the lowest R^* in E120 – *Andropogon gerardi* and *Schizachyrium scoparium* – grown along a soil fertility gradient in E055 (Dybzinski & Tilman 2007). Though mean trait estimates suggest that *S. scoparium* is a superior nitrogen competitor, results from E055 suggest that *A. gerardi* is a superior competitor in soils with greater than 0.1% total soil nitrogen concentration (Fig. S5A). While this is well above the initial soil nitrogen concentrations in E120, soil fertility has increased substantially over time because of feedbacks between plants and soils, particularly in diverse plots (Fornara & Tilman 2008). To account for this, we switched competitive hierarchy between *A. gerardi* and *S. scoparium* in 80% of simulations for planted richness treatments of 8 or more species. To better address these changes in competitive hierarchy in the model, we would need to know how soil fertility

gradients influenced competitive hierarchies for all species in the multi-species mixtures, and would need methods for predicting total soil nitrogen dynamics through time.

To test for the effects of including other forms of resource competition among legumes, we compared legume species abundances as a function of the abundance of other legume species in E120. Based on these comparisons, it appeared that the legume *Lupinus perennis* most strongly suppressed the abundance of other legumes (Fig. S5D). Though *L. perennis* has the highest R^* of all species in E120, this is not necessarily surprising, as legumes can fix their own nitrogen, and soil nitrate concentration in monoculture may therefore be an indication of fixation rate, rather than competitive hierarchy. To account for this, we switched the R^*-based competitive hierarchy of *L. perennis* with that for *Amorpha canescens* (the legume species with the lowest R^*) in 80% of simulations. To better address legume competition in the model, we would need more data on legume competition for other potential limiting resources, such as water, phosphorus, or light (Ritchie & Tilman 1995).

Finally, to test for the effects of seasonal changes in species traits, we augmented our model to allow *Poa pratensis* to access one third of its total nitrogen before it encountered competitive interactions with other species. This was meant to account for seasonal differences and rooting depth differences with may allow *P. pratensis* to avoid competitive interactions with many species that have lower R^* (McKane, Grigal & Russelle 1990; Fargione & Tilman 2005). For example, *P. pratensis* grows most of its biomass early in the season, before more dominant competitors, such as the forb *Liatris aspera*, reach their peak biomass (Fig. S5G). To better address seasonality in the model, we would need to account for covariance in traits among species across space and time (i.e. to determine how competition among species changes by season and by soil depth), but this would require replicated data from monocultures growing in the same locations during the same years, which we lack for this study.

Appendix XII. Source code for replicating study

The full source code for replicating the analyses in this manuscript, including all data needed for fitting and testing models, is available at github.com/adamtclark/tradeoff_model

Supplementary Figures:

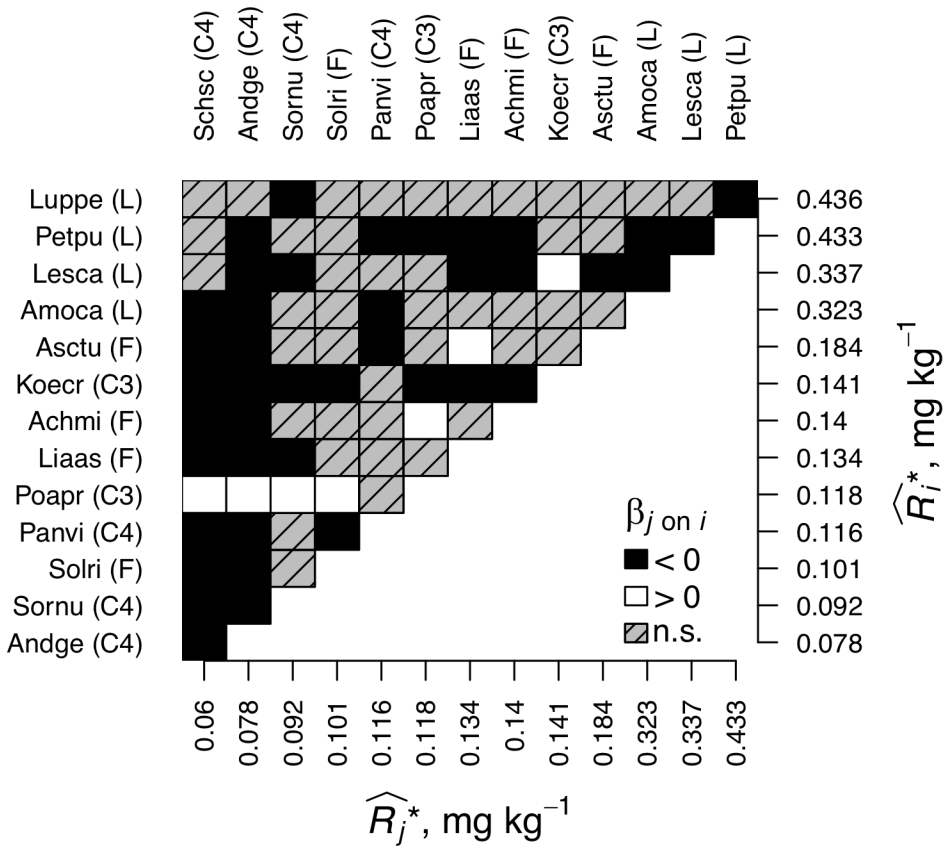

Figure S1: Pairwise relationships between species observed in experimental multi-species communities in E120. $B_{j \text{ on } i}$ describes the statistical effect of superior competitor species *j* (columns) on inferior competitor species *i* (rows), with competitive hierarchy corresponding to $\hat{R}^*$. Letters in parentheses correspond to functional groups: C3/C4 grass, (F)orb, or (L)egume.

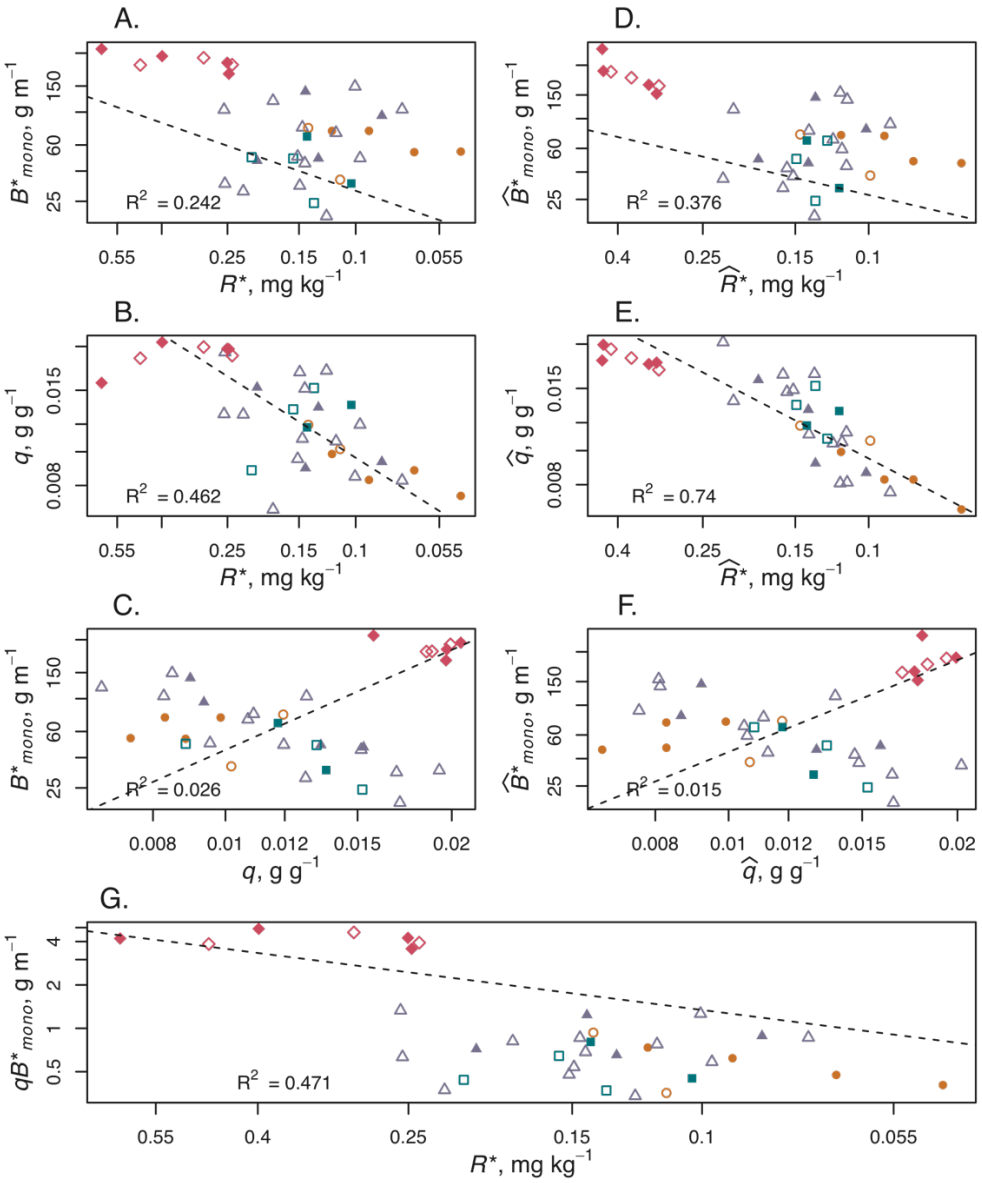

Figure S2: Bivariate fits between all pairs of traits used in the analyses. **(A-C)** show pairwise comparisons for raw trait values, and **(D-F)** show snapped traits. **(G)** shows bivariate relationship between the two major groupings of plant traits that we use in our model, qB^*_{mono} and R^*. Dashed line and R^2 values correspond to ranged major axis regression. Note that simple pairwise relationships explain much less of the variation than the three-dimensional fitted plane, even for snapped traits. Symbols and colours show functional groups and identify species planted in the multi-species plots, as described in the legend for Fig. 2 in the main text.

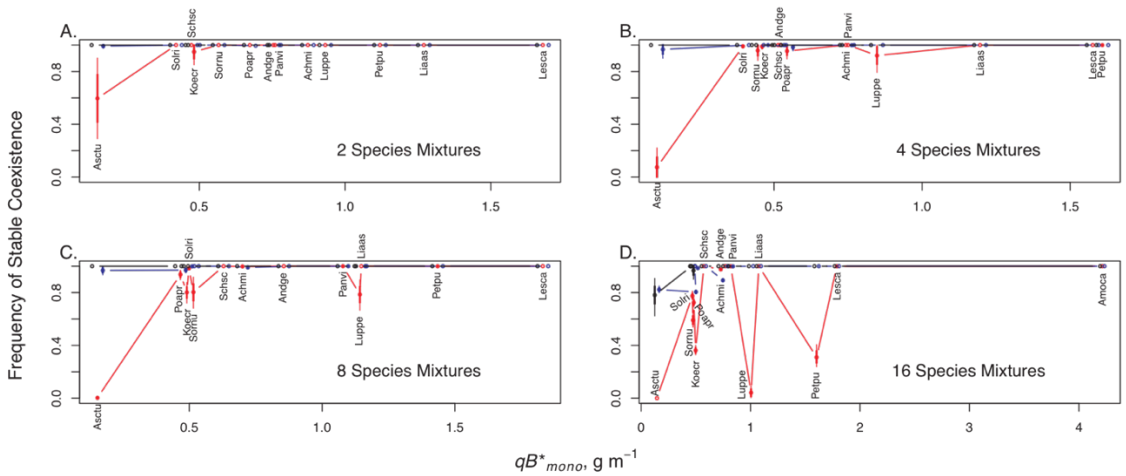

Figure S3: Coexistence in observed communities and model predictions, as a function of species mean monoculture aboveground tissue nitrogen. Intervals show mean ± one standard deviation, and 95% confidence interval. Black lines show fraction of multi-species plots in which planted species have persisted. Red and blue lines show results for mechanism-based models parameterized with raw and snapped traits, respectively. Stable coexistence in these models is identified using linearized stability analysis, as described in the methods section of the main text. Note that most species appear to stably coexist in the observed plots, as predicted by the model parameterized with snapped traits. For the model parameterized with raw traits, species that have both low qB^*_{mono} and low R^* (e.g. *Achilliea millefolium*) tend to be competitively excluded.

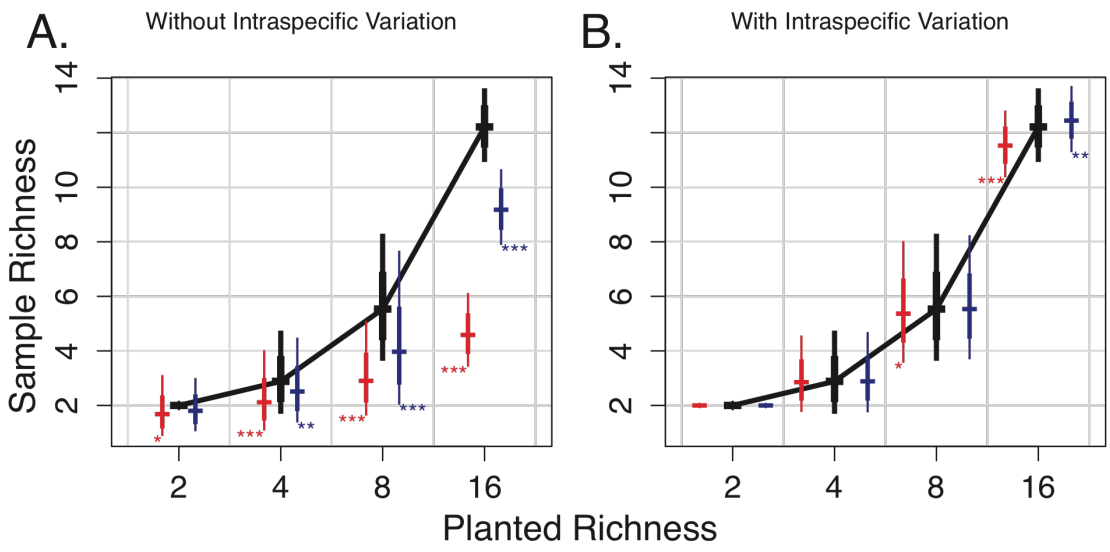

Figure S4: Observed and predicted species richness across diversity treatments for models with and without intraspecific trait variation. Intervals show 95% confidence interval and mean ± standard deviation. Black, red, and blue intervals denotes observed richness, and fits for raw and snapped traits, respectively. Asterisks signify significant differences between observations and predictions (*p<0.05; **p<0.01; ***p<0.001); p-values are from two-sample Wilcoxon tests.

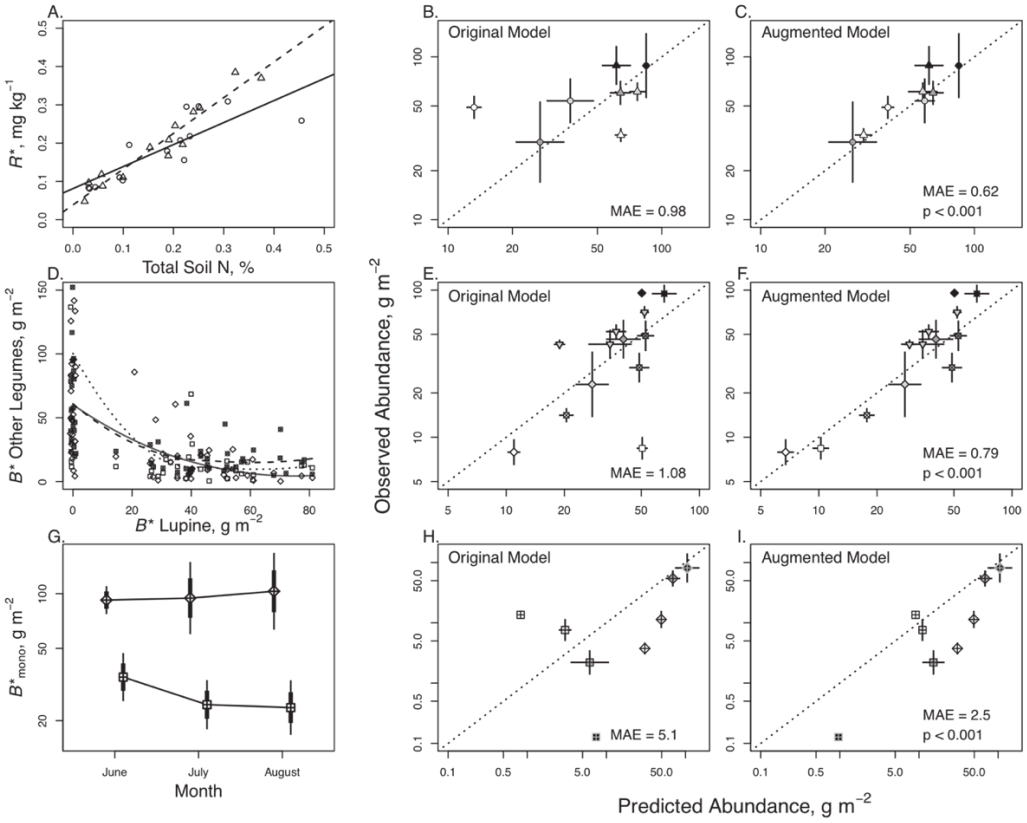

Figure S5: Examples of model augmentations. Accounting for these factors significantly improves model fit. Intervals show mean ± standard error for observed and predicted biomass across all plots in each planted richness treatment. P-values show significant differences in MAE between original and augmented models from two-sample Wilcoxon tests. **(A-C)** *A. gerardi* (circles/solid line) and *S. scoparium* (triangles/dashed line) switch competitive hierarchy in rich soils. **(D-F)** Though it has a higher *R**, *L. perennis* (triangles) appears to suppress the legumes *A. canescens* (boxes/dotted line), *Lespedeza capitata* (crossed-out circles/dashed line), and *Petalostemum purpureum* (diamonds/solid line) (values of zero biomass for *L. perennis* are jittered for clarity). **(G-I)** Early-season species such as the cool-season grass *P. pratensis* (boxed triangles) grow the majority of their biomass early in the year, while more dominant nitrogen competitors such as the forb *Liatris aspera* (gridded boxes) grow much of their biomass late in the year. Allowing *P. pratensis* to access resources before other species in one third of diverse plots leads to increases in its abundance, and decreases in warm season species abundances. Lines in **(G)** show mean ± standard error and 95% confidence interval.